RPWG
JJ

MONITORING RECOVERY FOLLOWING THE *EXXON VALDEZ* OIL SPILL: A CONCEPTUAL MONITORING PLAN

DRAFT FINAL

Prepared for

RESTORATION PLANNING WORK GROUP
Exxon Valdez Oil Spill Group
645 "G" Street
Anchorage, Alaska 99501

Prepared by

PARAMETRIX, INC.
5808 Lake Washington Blvd. NE
Kirkland, Washington 98033

In association with

ABA Consulting
PO Box 1151
Capitola, California 95010

Goldstream Consulting
PO Box 83090
Fairbanks, Alaska 99708

Darryll Olsen
2980 Glenmorrie Drive
Lake Oswego, Oregon 97034

May 18, 1993

TABLE OF CONTENTS

Page

TABLE OF CONTENTS
(continued)

APPENDICES

LIST OF TABLES

LIST OF FIGURES

SUMMARY

Monitoring of injured resources and damaged services is necessary to document when and if recovery occurs in response to the 1989 *Exxon Valdez* oil spill. Recovery monitoring will provide information on:

- Natural and assisted recovery
- The effectiveness of restoration activities
- Identification of the need for additional restoration activities
- General health of the ecosystem to better understand and respond to future perturbations

Additionally, recovery monitoring may provide information on sublethal effects from the spill and identify areas warranting research.

This report constitutes completion of Phase 1 of a three phase process to develop a recovery monitoring program:

- Phase 1 is the development of a conceptual monitoring plan. The conceptual monitoring plan provides a framework for Phase 2, by:

    - Providing examples of conceptual models from which to build resource- and service-specific models

    - Outlining and prioritizing the needs and objectives of the monitoring, and the strategies to meet the needs

    - Identifying recovery endpoints

    - Providing a mechanism for prioritizing monitoring activities

- Base on the framework presented herein, during Phase 2 the plan can be expanded to detail resource- and service-specific monitoring components (such as select endpoints), determine frequency of monitoring, geographic area to monitor, statistical methodologies, linkages in the ecosystem, etc.

- Phase 3 of the monitoring is actual implementation. at this stage contractor(s) will be awarded contracts to monitor recovery of injured resources and damaged services.

Development of Phase 2 and implementation of the monitoring in Phase 3, depend, in part, on which of the five restoration alternatives presented in the draft Restoration Plan is selected. The scope of the monitoring and restoration research varies with each restoration alternative based on the allocation of funds for monitoring.

The conceptual plan has been developed with the assistance of a diverse group of individuals, through implementation of a telephone interview process of over 50 individuals, and through presentation of a three-day workshop to discuss key issues and test mechanisms for prioritizing monitoring activities. Participants in the process, other than the project team, included members of the Restoration Team (RT), Restoration Planning Work Group (RPWG), approximately 35 experts/peer reviewers, approximately 35 principal investigators that participated in the Natural Resource Damage Assessment and/or Restoration Science studies, staff from the Regional Citizens Advisory Council for Cook Inlet and for Prince William Sound, and additional agency staff. This document represents a synthesis of the views and ideas of these people.

The conceptual plan is compiled in nine sections, as follows:

- Section 1 includes a general overview of the program, why monitoring is important, and the use of a conceptual plan.

- Section 2 includes a discussion of the value of and constraints on monitoring.

- Section 3 includes the definition of recovery and the various monitoring elements.

- Section 4 presents the needs and objectives, along with the strategies to address them.

- Section 5 presents potential recovery endpoints for the injured resources and damaged services, a proposal for development of resource- and service-specific conceptual models, and a mechanism of prioritizing monitoring activities through the application of criteria and a ranking system. The results of a trial application of the criteria are presented with recommendations on how to improve upon the trial.

- In Section 6 general guidance on sampling design is provided including information on methodologies for monitoring, focusing primarily on statistical elements.

- Section 7 presents ideas for implementation of the monitoring.

- Section 8 includes recommendations for Phase 2.

- Section 9 presents references cited.

The primary elements of the framework presented herein, include: (1) the recommendation of mechanisms for prioritizing monitoring activities, including the development of conceptual models, (which should be developed on a resource- and service-specific basis in Phase 2); (2) the development and application of criteria for addressing (1); (3) the development and prioritization of needs and objectives of monitoring; and (4) identification of the relationships between resources and services and between the *Exxon Valdez* monitoring program and other programs within the spill area and outside Alaska. Results from this process provide information useful to development of a consistent, comprehensive program in Phase 2.

Consensus building is a key component of the conceptual monitoring plan. Consensus building has been sought through both the numerous interviews and the workshop. Consensus building should continue through Phase 2 of this project to provide maximum acceptance of the results.

1. INTRODUCTION

1.1 BACKGROUND

The *Exxon Valdez* Oil Spill Trustee Council (herein referred to as the Trustee Council) is developing a Restoration Plan for the spill injured resources and damaged services. One option under consideration during development of the Restoration Plan is to implement a comprehensive monitoring program to monitor recovery. This report addresses Phase 1 of the monitoring program, the development of a conceptual monitoring plan. The purpose of the monitoring program is:

- To assess the adequacy or effectiveness of both natural and assisted recovery
- To document long-term trends in the condition of resources and services affected by the oil spill
- To contribute to existing physical, chemical, and biological baseline data on resources and services in the spill area

The Trustee Council initiated a planning effort to develop the first phase of a comprehensive and integrated monitoring program for resources injured and services damaged by the 1989 *Exxon Valdez* oil spill. The conceptual monitoring plan (Phase 1) will provide the framework for the more detailed technical planning during Phase 2, followed by implementation in Phase 3. The framework will be used by the Trustee Council to make decisions involving the selection and implementation of monitoring activities. The conceptual monitoring plan resulting from Phase 1, or elements thereof, will be incorporated into the Restoration Plan. Phase 1 recommends mechanisms for prioritizing monitoring activities, sets goals and objectives for monitoring, identifies relationships between monitoring components, and identifies existing monitoring programs and potential linkages. The intent of the conceptual plan is to provide objectivity to the decision-making process. During Phase 2 the framework will be expanded and refined to include resource- or service-specific programs and methodologies, including development and review of conceptual models, sampling designs and statistical approaches. The conceptual models developed in Phase 2 will be applied to the injured resources and damaged services to ensure proper feedback mechanisms exist to determine if the goals and objectives are being met.

1.2 WHY MONITOR RECOVERY?

The question, "why monitor recovery?", requires a two part answer. First, monitoring is key to determining if recovery has occurred. The rate of recovery of resources and services can be established through monitoring, providing insight as to which resources and services may need assistance to recover. However, recovery of resources and services is not only a function of whether or not they have reached pre-spill conditions, but also a function of the public's perception and use patterns of those resources and services. This perception can only be based on reality if monitoring occurs. Likewise, decisions in managing the resources and services are largely influenced by the public's perception of resource and service recovery. These perceptions should also be based on information that can only be derived from monitoring recovery.

The second part of the answer to this question is that the credibility of the Trustee Council in making decisions regarding recovery also requires monitoring. The general public, special interest groups (e.g., subsistence, commercial fisherman), and agency technical staff cannot be expected to support decisions of the Trustee Council in the absence of data documenting the status of resources and services.

Thus, monitoring is an essential component of documenting recovery. Only through an adequate degree and duration of monitoring can the Trustees fulfill their responsibility to provide stewardship in the recovery of the injured resources and damaged services.

1.3 WHAT IS A CONCEPTUAL MONITORING PLAN?

A conceptual monitoring plan is an instrument identified by the National Research Council (1990) in *Managing Troubled Waters* as a means to logically direct our nation's environmental monitoring. Its ultimate goal is to guide the planning and decision making process in any monitoring program to produce information that is useful in making management decisions and to communicate the status of natural resources to various interest groups. To reach this goal there must be considerable two-way communication between scientists generating information and users of the information (management agencies and public).

The National Research Council describes a conceptual monitoring plan as:

- A tool for developing and refining monitoring systems
- A means for identifying elements to be considered for an optimum monitoring plan
- A guide for decisions on what to monitor, when, how, and where

A conceptual monitoring plan is a means for establishing a relationship between those who require monitoring information and those who provide this information. It is a generic plan for establishing criteria and procedures desirable for implementing specific monitoring plans. It is a guide to decision making regarding monitoring activities. It provides guidance in dealing with variability and uncertainty in monitoring. The plan also provides a map for coordinating various monitoring activities.

As with any such tool, it is both how well the tool is constructed and how well the tool is used that determines its effectiveness. Our basic precept in constructing this conceptual monitoring plan is that it be the product of contributions by as many involved parties as possible. Thus, we have actively sought the participation of a large number of individuals through telephone interviews, a technical workshop, and by review of previously prepared materials.

1.3.1 Monitoring Plan Principles

There are two basic principles inherent in the conceptual monitoring plan. These principles follow:

- Whenever possible, monitoring designs should reflect cause-effect relationships while accounting for variability and uncertainty.

- Specific design decisions (e.g. the number of stations, number of replicates, monitoring procedures, etc.) can be made only after objectives and related information needs are clearly established.

The goal of producing information that is useful in making management decisions will only be met if these basic principles are followed.

1.3.2 Essential Elements of a Conceptual Monitoring Plan

There are a number of elements essential to a conceptual monitoring plan as identified by the National Research Council (Figure 1). These elements include:

Needs:	To be successful, a conceptual monitoring plan must take its direction from the needs of the eventual users of the information produced by the plan.
Users:	Those who require monitoring information for management or
Environmental Conditions:	Knowledge of the existing basic features of the environmental resources and services these resources support.
Objectives:	Clear statements of the needs and expectations the users have for the monitoring program.
Investigators:	Those who will develop and implement specific monitoring plans, analyze results, and communicate monitoring information.
Sampling Design:	Technical approach for the hypotheses to be tested; what, how, where, and when to monitor; and how data will be analyzed.
Implementation:	Strategy for establishing and maintaining monitoring activities and communicating information.

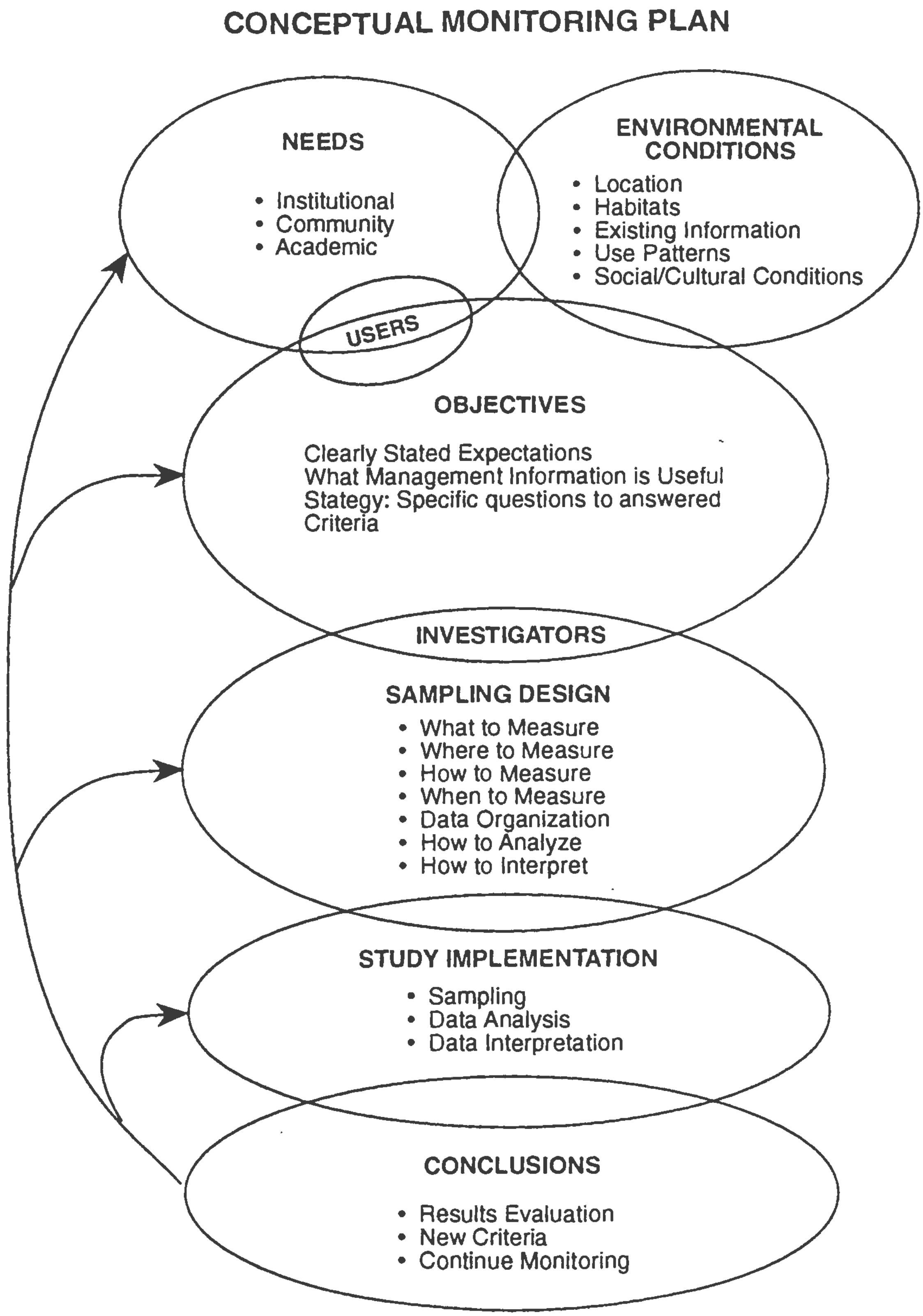


Figure 1.
Essential Elements of a Conceptual Monitoring Plan

Evaluation: Evaluation of the results and conclusions as a feedback mechanism to assess whether monitoring has been effective at documenting recovery, and whether or not monitoring should be continued.

How these elements flow from the development of goals (needs) and objectives through actual implementation of monitoring is illustrated in Figure 2.

1.3.3 Needs and Expectations

The monitoring needs and expectations, biology of the resources, and characteristics of the services will define what information [objective(s)] is useful to the Trustee Council and investigators attempting to determine when resources and services have recovered or at what rate they are recovering. Development of the objectives requires communication between the users of monitoring information and the investigators, designers, and analysts who will produce this information. Development of the objectives also requires integration of public concerns and expectations together with the legal framework (Settlement Agreement).

These objectives should be unambiguous statements defining what constitutes useful information. They should require a cumulative assessment approach to provide a synoptic view of the injured resources and damaged services. This synoptic view should:

- Identify the recovery of multiple resources and services as well as cumulative recovery of the ecosystem.

- Describe levels of certainty anticipated in recovery endpoints (definition of variation)

- Provide a framework for synthesizing monitoring information

1.3.4 Study Strategy

The objective of developing a study strategy is to narrow the focus of monitoring efforts on questions and parameters of the resources and services that are most likely to produce the needed information. The study strategy identifies the resources (species) and services at risk or sufficiently in need of recovery monitoring. It also involves development of conceptual model (not conceptual plan) that clearly state questions able to be tested.

Figure 3 illustrates the basic elements of such a conceptual model (not conceptual plan) for recovery monitoring. It illustrates that the Trustee Council, together with the investigators and interested public, should be involved in developing expectations. This conceptual monitoring plan involves the development of Trustee Council and investigator expectations. The plan indirectly includes public participation through the Public Advisory Group's review and comments provided to the Trustee Council. This participation has lead to the development of the goals and objectives of this conceptual monitoring plan.

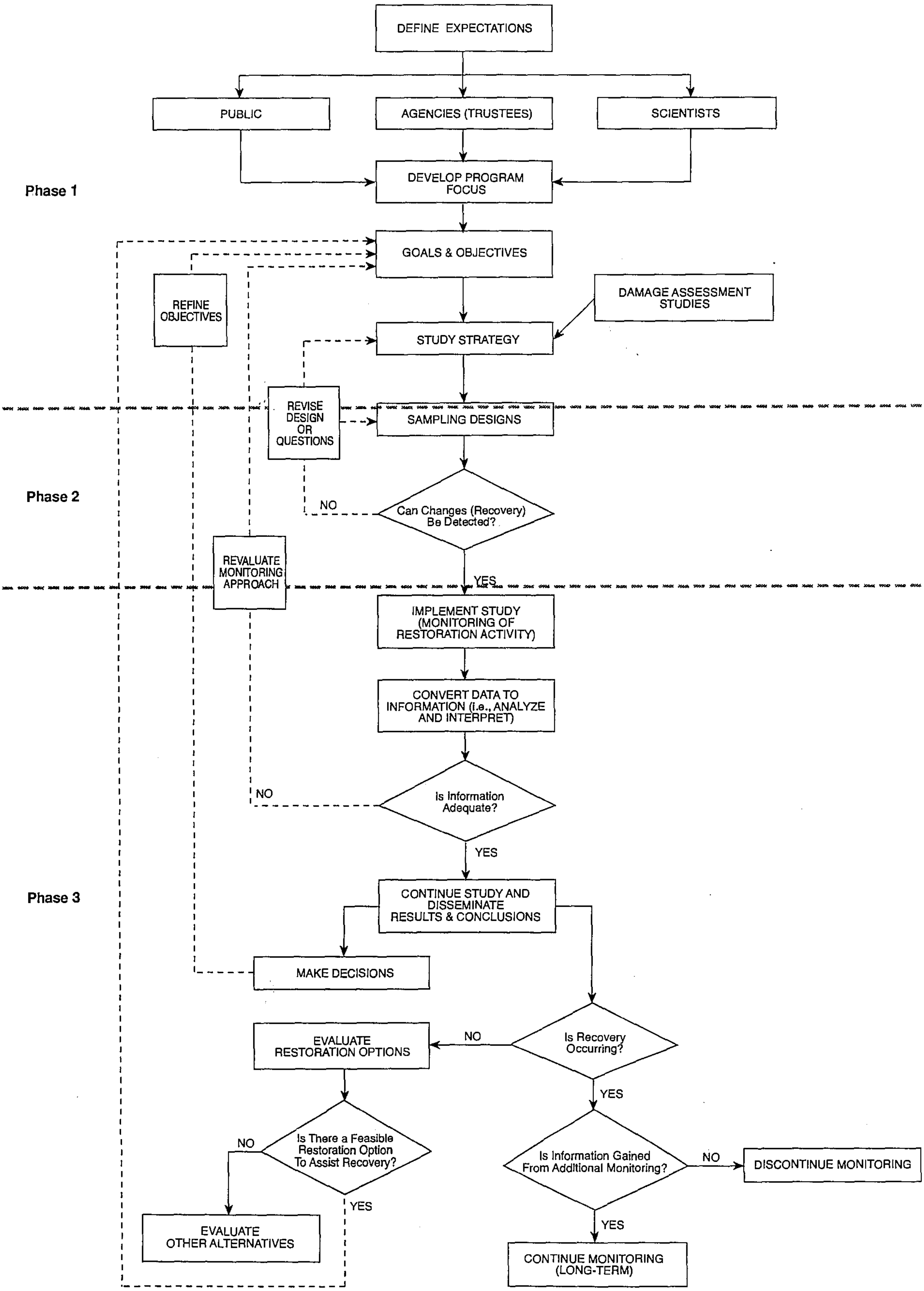
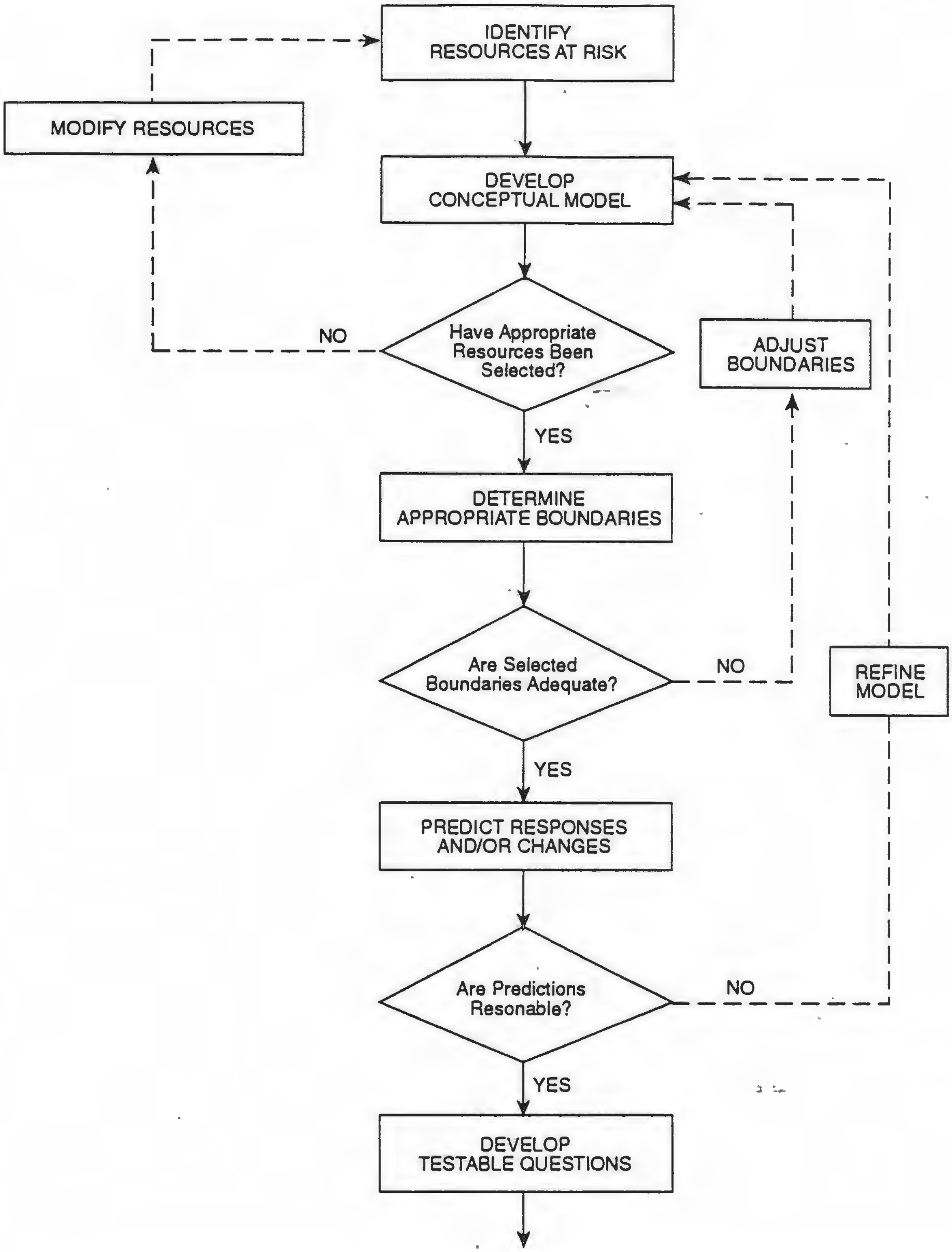


Figure 2.
Design and Implementation Elements, Recovery Monitoring Studies



Source: NRC 1990.

Figure 3.
Defining Study Strategy

Participation by the various parties has also led to the formation of general strategies to reach the identified goals and objectives. These strategies attempt to:

- Define conceptual models
- Establish boundaries
- Develop predictions with estimates of uncertainty

The conceptual models include questions that are able to be tested. Clearly stated questions are a part of developing specific monitoring plans. These questions should identify links among ecosystem attributes. The questions must be able to be tested within the constraints of the ecosystem, scientific and survey techniques, and financial resources, as well as institutional constraints.

Boundaries established in the conceptual models include spatial, temporal, biological, physical and chemical boundaries, as well as social, cultural, and economic boundaries. These boundaries are based on information derived from the damage assessment investigations and restoration activities. Additional boundaries for the recovery monitoring are the legal constraints imposed by the civil Settlement Agreement and the practical, but undefined, boundary established by available funding.

Conceptual models should identify quantitative and qualitative changes in the resources and services expected during recovery. These predictions should attempt to identify the effects of resource and service management actions on the targeted resources and services. They should also identify the uncertainty likely to exist in measuring or estimating these changes.

Finally, the conceptual models should provide for review of predictions and questions that can be tested during the coarse of investigations. These reviews should lead to refinement and reformulation where appropriate.

1.3.5 Preliminary Studies

The conceptual monitoring plan developed by the National Research Council (1990) identifies preliminary research as a key step to developing specific questions. In the case of the *Exxon Valdez* oil spill, the majority of preliminary studies have taken place in the form of natural resource damage assessment (NRDA) studies (see Figure 2). These studies are often more extensive, in both areal extent and degree of investigation, than is likely to be considered in most preliminary studies. The completed NRDA studies are generally adequate to fill the role of preliminary investigations for the purposes of recovery monitoring, although not all have been completed or made available to the users. In addition to the NRDA studies, other monitoring studies undertaken by resource agencies provide some information that could also fill the role of preliminary investigations.

1.3.6 Sampling Design

The key component in the sampling design for specific studies is the link to questions that are able to be tested. Many of our nation's past monitoring studies have failed to meet expectations because they failed to link monitoring to questions that can be tested. It is important that monitoring projects explicitly state what they intend to accomplish and be held accountable for accomplishing specific objectives. Other key elements of a sampling design are discussed below. Information on development of sampling designs for specific monitoring activities in the oil spill area will be prepared as part of Phase 2 of the monitoring program. It is recommended that conceptual models be used to aid this process.

Sampling designs have a number of key elements, most of which are obvious to investigators. The key elements were reviewed by the National Research Council (1990) to ensure that they are included in any well-planned monitoring activity. These key elements are:

- Identification of meaningful kinds and amounts of change (e.g., time/spatial scales)
- Identification and quantification of sources of variability, both natural and within the sampling design
- Specification of how variability will be partitioned
- Decisions of what to measure (which will require knowledge of resource biology and population dynamics and characteristics of services).
- Statistical models for selection of kinds and numbers of measurements
- Optimization and power analyses to ensure detection of meaningful levels of change
- Quality assurance objectives

Meaningful change is based on the questions derived in the preceding step. Both the users and the investigators must contribute to defining what types and levels of change can be measured and how they will identify recovery of the resource and service.

There are many sources of natural variability that should be considered in developing a monitoring plan. Seasonal, cyclic, successional, and biological interactions, as well as cultural and human use interactions, are major sources of variability to be considered by investigators in developing specific monitoring plans. Although characterizing variability is difficult because of the many sources, or incomplete understanding of the sources, and the scale of the marine environment, it is essential to develop a capacity to detect meaningful levels of change. Other man-caused sources of variability should also be considered.

Variables selected to measure recovery of resources and services should focus on those most likely to provide information on recovery. The variables should address the questions that can be tested and are identified in the preliminary studies (e.g., NRDA studies and preliminary restoration studies). The choice of variables should be based on knowledge obtained through review of these preliminary studies, and on the statistical capability to reflect change. Variables can include:

- Early warning indicators (those most likely to detect recovery)
- Sensitive indicators (those which have had the greatest damage)
- Process indicators (those reflecting complex system interactions)
- High information indicators (those representing a number of different parameters or resources and/or services)

Statistical models should identify the more precise estimates that can be derived for the smallest sampling effort. This will help to select variables with information-to-noise ratios that are adequate to test the identified questions. The statistical models should define how questions will be evaluated and how variations from other sources will be interpreted. Statistical comparisons should be evaluated to consider the capacity to compare baseline conditions or reference areas, both of which are commonly difficult.

Sampling optimization and power analyses are means of both ensuring that objectives are met and that appropriate levels of effort are employed. These statistical techniques require quantitative estimates of the major sources of anticipated variability. Their application should lead to appropriate allocation of limited financial resources.

Quality assurance activities include quality control and quality assessment. Quality control is included within specific monitoring plans to ensure standardization of sample collection, processing protocols, analytical techniques, and technician training. It should provide a means to correct or remove erroneous data and resolve inconsistencies that degrade data integrity.

Quality assessment requirements are also incorporated into specific monitoring plans. They quantify the effectiveness of quality control procedures by instituting repetitive measurements, internal test samples, interchange of operators and equipment, independent verification of findings, and audits.

To be effective, quality assurance must be included in initial planning of the monitoring program. It must continue as an integral component of the total monitoring system through implementation and dissemination of information.

1.3.7 Data Conversion to Information

The objective of monitoring is to produce useful information rather than volumes of data. Through the organization, processing, and synthesis of data, together with knowledge, the data are endowed with reference and purpose, to become useful information. This useful information provides additional knowledge to be used in making decisions. Conversion of data to useful information involves planned data management, as well as planned data analysis and modeling.

Data management should be planned to provide easy access to data and related information by all users, including resource and service managers. Because of the amount, complexity and inter-relationships of data, it is essential to establish a computer-assisted data management system. It may be that all data should be stored in a central location or library. The data management system should consider data quantities, relationships of various data, quality assurance requirements, and types of analyses to be performed. "Data management activities are as important to the success of monitoring programs as the collection of data." (National Research Council 1990).

The objective of data analysis is to summarize and simplify data, to test hypotheses, and to measure change (recovery). These analyses should be planned as part of development of specific monitoring plans. To be successful, the analyses should summarize results, deal with linkages among data, use standard modeling approaches, evaluate assumptions, and evaluate sensitivity of analyses.

1.3.8 Dissemination of Results and Conclusions

It is obviously important that results and conclusions be disseminated to the users. Mechanisms and timing of reports to accomplish dissemination should be included in development of the monitoring plans by requiring this in the proposal and contracting process. Status reports should be included to allow evaluation of monitoring efforts and adjustments where appropriate. Management information is only produced when it is actually conveyed in a usable and accessible form.

1.4 MONITORING PLAN APPROACH AND DESIGN

Development of this conceptual monitoring plan relied, in part, on the report, *Managing Troubled Waters, The Role of Marine Environmental Monitoring,* produced by the National Research Council in 1990. This report describes the role of a conceptual monitoring plan in guiding monitoring efforts and provides guidance in preparation of a conceptual plan. This plan also relies heavily on the input and advice from resource and service experts, principal investigators, agency representatives, and Restoration Team and RPWG members. The various components of the conceptual plan are, in large part, a synthesis of ideas and contributions obtained by interviewing these individuals, and through a three-day workshop which many of them attended. The value of the conceptual monitoring plan is derived primarily by their contribution.

Development of the conceptual monitoring plan began with of development of a preliminary draft of the plan through an interview process and workshop format. A questionnaire was developed (see Appendix A) to establish what user groups expected from the conceptual monitoring plan. Approximately fifty individuals were queried, (see Appendix A for the list of interviewees), including Restoration Team and Restoration Planning Work Group members, peer reviewers and principal investigators. These interviews were synthesized to form the draft plan. To aid in the refinement of the preliminary draft plan, a three-day workshop was held in Anchorage, Alaska to confirm the following:

Day One: Establish that the intent of the RPWG was met with the development of the preliminary draft plan and workshop format, and to receive RPWGs comments for development of the final plan.

Day Two: Conduct key informant interviews with peer reviewers/experts to address the following key issues:

- Identification of what constitutes recovery.
- Prioritization of needs and objectives.
- Determine if strategies address the objectives.
- Identify the strengths and weaknesses of the proposed data management network.
- Development of a mechanism for setting monitoring priorities.
- Discussion on the management structure for monitoring.
- Review of other monitoring programs.
- Identification of stresses known to cause population effects.

In Day One of the workshop it was determined that issues one, two and five were of the greatest concern, thus the primary focus on the key informant interviews (Day Two workshop activity) was on these items.

Concurrent with the key informant interviews was a brain-storming session of the peer reviewers (who were not being interviewed at the time) to apply the criteria for prioritizing monitoring activities, and to review other monitoring programs in light of the recovery monitoring program.

Day Three: Provide a working forum with participation of those initially interviewed plus other interested parties. The goal was to reach agreement on the overall needs of the monitoring program, to establish a mechanism for prioritizing monitoring activities, and to reach agreement on the criteria to be used in evaluating monitoring activities and for establishing recovery endpoints for injured resources and damaged services.

Recovery of all of the resources injured and services damaged or potentially injured/damaged by the *Exxon Valdez* oil spill cannot be monitored equally in time and space. The approach employed by this conceptual plan is to design a recovery monitoring program that accomplishes the following:

- Defines the goals, objectives, and strategies for monitoring recovery

- Identifies resources and services, or elements thereof, that should be considered for monitoring

- Provides a mechanism to monitor on an annual basis and throughout the life of restoration funding (i.e., 10 years) and beyond

- Provides guidance on considerations of sampling design

- Identifies opportunities for comparing monitoring activities to and integrating them with other programs

- Identifies potential mechanisms to guide implementation of the recovery monitoring program

A primary objective of the conceptual monitoring plan is to serve as a tool in designing specific monitoring plans for the resources and services to be monitored. Figure 3 shows a conceptual model of how this tool can be applied to recovery monitoring in the oil spill area. In order to narrow the focus of monitoring efforts, a model should be developed for each resource and service that addresses the elements in Figure 3. Setting the boundaries around the questions such that they are testable hypotheses and answered is a critical element.

Additionally, the conceptual monitoring plan specifically addresses the questions stated in the contract scope of work for the development of plan. These are stated below with a summary response. Throughout the plan each question is discussed in more detail.

1. What process or mechanism would best assist the Trustee Council in determining monitoring priorities?

There are several mechanisms that could be used for prioritizing monitoring activities; however, those we recommend are:

- Consensus building through participation of the various user groups
- Development of recovery endpoints
- Development and application of criteria for evaluating monitoring activities
- Development and application of conceptual models for injured resources and damaged services

These are further discussed below.

Consensus Building. As stated in the civil settlement, public involvement is to be an integral part of the restoration process. In order for the public to feel that recovery has been successful and that the settlement monies have been used properly, they need to be involved in the process. Their input will help determine acceptable monitoring endpoints. In order to gain the maximum knowledge and perform a scientifically credible program, resource and service experts must also be involved throughout program development, implementation, and review. Thus, an important element in determining monitoring priorities is the involvement, review, and consensus by the various user groups.

Recovery Endpoints. The civil settlement from the *Exxon Valdez* oil spill resulted in money set aside specifically to restore, replace, enhance, or acquire equivalent resources that were injured due to the spill, and the reduced or lost services. The prioritization of activities funded by the monies for monitoring should be driven by monitoring and recovery endpoints and public concerns. Specific recovery endpoints for a particular resource or service should be developed, as should monitoring endpoints. The difference between the two is that some resources or services may be monitored beyond the defined recovery endpoint (i.e., long-term monitoring). Criteria should be developed to assist the Trustee Council in determining when to continue monitoring beyond recovery. Lastly, endpoints for long-term monitoring should be developed (which may include some of those for continued recovery monitoring). Development of endpoints is necessary because long-term monitoring that answers questions (beyond recovery endpoints) about an ecological or social interaction may not be useful once the mechanisms are clearly identified.

Development and Application of Criteria. It will be useful to know how the public would like to see monitoring monies spent and what resources and services they are most concerned about. The socioeconomic value of a monitoring action can be related to what the public are willing to pay for monitoring and/or restoration. This input may be ascertained by the public's review of the draft Restoration Plan.

To gain a scientific perspective, the resources and services and potential monitoring activities should be prioritized to determine what activities will provide the most information. This should then be compared to the costs, and potentially reordered slightly to gain the most information for the money. Finally, the public feedback and scientific perspectives must be integrated. If, for instance, the public feels that monitoring of killer whales is important, this activity must be compared to the prioritization of monitoring other injured resources and damaged services, to determine the benefit (both scientific and public perception) from such a monitoring activity.

To further address the scientific perspective of the prioritization of monitoring activities, a matrix can be used to assess linkages between the resources and services and between potential monitoring activities. In addition to prioritizing overall monitoring activities, it is necessary to prioritize activities specific to a resource or service, an activity recommended in Phase 2. Priority should be given to activities that are most likely to address the needs and objectives of recovery monitoring.

The workshop conducted during in development of this plan involved applying a ranking system to prioritize the resources and services to be monitored. The process involved reaching agreement on the criteria on which to evaluate resources and services to monitor, then applying a rank to each criteria for each resource and service (see Section 5.4). This process appears somewhat successful in prioritizing the resources and services to monitor, but further refinement is recommended as discussed in Sections 5 and 8.

Conceptual Model Development. Conceptual models are a method to illustrate the links between resources, services, ecosystem processes, and anthropogenic influences. They can be used as a tool to assist in deciding monitoring priorities based on resource and service interactions. Conceptual models help define cause-and-effect relationships or expected relationships, and permit hypotheses to be developed, as well as assist in development of specific monitoring strategies.

Conceptual models can include many types of information, such as natural history information, subjective judgement, ecological theory and numerical models (National Research Council 1990). The information gathered through the NRDA, restoration activities, and from the literature, and activities such as the workshop conducted as part of this program, can be factored into conceptual models for each resource and service. Through the interviews, the workshop, and from the development of this plan, the linkages between resources and services, as well as the identification of recovery endpoints, were discussed — both of which are important to developing conceptual models. Additionally, the identification of needs, objectives and strategies of the monitoring plan as identified in Section 4, should be referred to in the development of the conceptual models in order to ensure that the overall monitoring objectives are met.

Another potential tool considered was the use of adaptive environmental assessment (see Environment Canada 1982). Although adaptive environmental assessment was considered, the primary tools discussed throughout this report are the development and application of conceptual models, consensus building through participation of all of the user groups, and the development through this process of criteria for prioritizing monitoring activities. However, several of the principles of adaptive environmental assessment are elements of the mechanisms described above.

2. What are realistic goals and objectives for monitoring?

Monitoring is essential to understanding if settlement activities have been successful at restoring, enhancing or replacing resources and services. The overall goal of the monitoring is stated in the draft Restoration Work Plan (*Exxon Valdez* Oil Spill Restoration Office 1993). The overall goal is to develop a comprehensive and integrated monitoring program that will:

- Follow the progress of natural and assisted recovery.

- Establish an ecological, social, and, cultural baseline from which future disturbances can be evaluated.

This goal has been further broken down into specific objectives in Section 4 of this conceptual monitoring plan. The objectives have been reviewed by many individuals including Restoration Team and Restoration Planning Work Group members, peer reviewers and principal investigators. The objectives in this plan reflect the input received from this review process. The objectives, as stated, are comprehensive and need to be further refined in Phase 2 of the monitoring program, when the bounds (financial, social, scientific, and political limits) of the monitoring program are set. The objectives stated in Section 4 are realistic in terms of technical feasibility and scientific merit. Cost ramifications, economic feasibility, and public acceptance of monitoring alternatives will not be determined until Phase 2 of this program. With each subsequent phase of the monitoring program, as well as during proposal review and throughout actual monitoring, the progress and specific elements should be reviewed to determine how well objectives are being met.

Since the extent of monitoring is monetarily driven, and the ultimate financial feasibility, rather than technical feasibility, scientific merit (including how well the program element addresses the objectives), and public concerns/interests, cannot be ascertained by those other than the Trustee Council. The Trustee Council will need to ultimately determine in which monitoring activities to invest, given the constraints of the available budget.

From the interviews conducted with peer reviewers, principal investigators, and Restoration Team and Restoration Planning Work Group members, the expectations for what the overall monitoring program should accomplish follow:

- Identify what recovery monitoring should and should not attempt to accomplish.

- Identify monitoring goals.

- Establish criteria for selecting resources and services to be monitored.

- Identify strategies to ensure recovery monitoring is effectively implemented.

- Ensure natural and sample variation is taken into consideration and provide guidelines.

- Identify and prioritize resources and services to be monitored and why.

- Identify appropriate monitoring approach (e.g., indicator species, population level, trophic level, ecosystem characteristics) and provide rationale.

- Provide mechanisms for integration with other monitoring and management activities.

- Define why there is a need for monitoring.
- Establish requirements for dissemination of information.
- Provide baseline data for assessing future perturbations.
- Define "recovery".
- Establish a plan or framework to guide long-term monitoring.
- Recommend a mechanism for managing recovery monitoring.

Most of these expectations are addressed in Section 4 by the stated objectives and strategies for meeting the objectives.

3. What resources and services should be monitored and why, given the goals and objectives of the monitoring?

A set of criteria were developed to assist in prioritizing the resources and services to be monitored, through a process of interviews and the workshop, with resource and service experts, principal investigators, and Restoration Team and Restoration Planning Work Group members (see Section 5). As part of the workshop activity, the criteria were classified as primary or secondary based on their relative significance in meeting the objectives of the monitoring program. A preliminary attempt at applying the criteria to the identified injured resources and damaged services was made during the workshop. The ranking was reviewed to establish prioritization of the resources and services to monitor, based on how well they met the criteria. Additionally, the criteria were applied to some resources not directly injured by the spill but identified as ecologically linked to the injured resources and damaged services, such as forage fish. The result of this process is described in Section 5 with further recommendations presented in Section 8.

The prioritization process described above only takes into account technical versus economic aspects of monitoring. During Phase 2 of the monitoring program, economic factors will be introduced within the technical design of each monitoring element (i.e., with proposal submittal for monitoring alternatives for each resource and service identified). The cost effectiveness of monitoring options as well as the application of the technical criteria will again be applied, this time to each monitoring option, to determine an overall prioritization of monitoring activities.

Additionally, during Phase 2, conceptual models will be developed for each resource and service, illustrating linkages, processes and known anthropogenic influences affecting each. These models will aid in prioritizing monitoring activities by assisting in developing testable hypotheses and establishing links between resources and services that may help with interpretation of monitoring results.

4. Which clean-up, damage assessment and restoration science studies contain elements that would best serve the purposes of the intended monitoring program, and what are these elements?

Damage assessment and restoration science studies that, to date, contain monitoring elements that address the overall goals (needs) of the monitoring program best serve the intent of the monitoring program. The programs that are continued or supplemented with monitoring, should remain consistent with the earlier studies (with standardized units of measurement, overlap of the parameters measured, and study of the same locations and populations, etc.) so that recovery is not measured differently than injury, and the data are useful in comparing to pre-spill or control area data.

Once the resources and services to be monitored have been prioritized (as described above), the clean-up, damage assessment, and restoration studies, and/or elements thereof, can be reviewed (during Phase 2) to determine which of these contain elements that would best serve the purposes of the monitoring program.

5. Which surveys of services (e.g., recreation subsistence, aesthetics, etc.) contain elements that would best serve the purposes of the intended monitoring program?

From the interview process two programs were identified as those that contain elements useful to the monitoring program: (1) The subsistence monitoring by Alaska Department of Fish and Game, which included both the monitoring of shellfish tissue concentrations, and of consumption levels; and (2) The sport and commercial fish catch data collected by the state and U.S. Fish and Wildlife Service. Of course, the usefulness of these or other surveys of services, may change according to the prioritization of resources and services for monitoring. However, both of the programs mentioned are the responsibility of resource management agencies, thus their continuation may not be dependent upon spill settlement funds.

Additionally, surveys of people's perceptions (i.e., key informant interviews, questionnaires) as well as evaluations of socioeconomic data associated with recovery of resources and services would be useful, since in the final outcome the public must feel that the activities funded by the settlement have yielded information on recovery of the injured resources and damaged services important to them. At least one such survey has been performed by RPWG members, a survey to assess the damages to services.

6. What consideration should be given to the relationships among different monitoring components (e.g., sediments, shellfish, fish, mammals, birds, etc.), and how should they be integrated?

Part of the overall goal of the monitoring plan is to follow the progress of natural recovery.

Because the recovery of one resource or service may be directly linked to the recovery of another resource or service, it is critical to look at the linkages between resources and services to measure and interpret recovery. Therefore, in evaluating monitoring priorities, the linkages between resources and services should be considered. That is why some of the criteria developed to aid in prioritizing monitoring of resources and services include linkages between the resources and services.

To facilitate review of the linkages between resources and services, a matrix has been developed of the damaged services and injured resources, including resources not directly affected by the spill but linked to the resources and services that were affected (e.g., mussels and forage fish). The damage assessment and restoration studies and results of interviews with experts and principal investigators were reviewed to construct the matrix. The matrix identifies relationships (both positive and negative) between resources and services and can be used as a tool to identify which recovery monitoring activities could be integrated. The matrix table is not considered to be complete. Feedback from resource experts is needed to develop a final matrix identifying linkages, and in some cases, additional research is needed on the natural history and processes.

It is equally important that the linkages between resources and services that are monitored be clearly presented in terms understandable to the public in order that they can understand the worth and value of the program.

As mentioned earlier, resource- and service-specific conceptual models will also aid in understanding and accounting for interactions that affect the methodologies and interpretation of monitoring results.

7. What relationships need to be established with other monitoring programs within the spill area and how should they be integrated?

There is value in identifying monitoring programs within and outside the spill area for several reasons:

- Monitoring may already be planned or underway that can provide answers to some of the objectives of this monitoring program, thus representing a savings of effort and money (e.g., Alaska Department of Fish and Game's (ADF&G) subsistence monitoring).
- Monitoring programs may provide information on methods, natural variation, and the usefulness of monitoring particular elements.
- Dove-tailing of programs may provide information on a global versus regional level.
- The monitoring in one program (e.g., effectiveness of restoration activity) may

management body of the monitoring program must work to organize interactive teams at the start of monitoring efforts and assure that these teams consider all the elements necessary to analyze recovery of resources and services of concern.

Depending on the funding level agreed to by the Trustee Council, monitoring may cover the two overall goals listed under 2 above. However, some monitoring of recovery and the documentation of long-term trends may fall outside the scope of restoration activities and as such, should be conducted by an independent party.

Monitoring the effectiveness of specific restoration activities as they relate to recovery should be a part of individual restoration activities. For instance, a restoration activity involving installation of a fish ladder that includes monitoring at the fish ladder to ascertain whether or not the ladder is effective in allowing fish passage and improving the rate of passage, should include a hypothesis of how increased fish passage is affecting recovery of the particular salmon run. However, recovery of an overall fisheries resource, regardless of location of the ladder, may be the responsibility of the recovery monitoring program. The two components (project and recovery) should be integrated so that the overall monitoring can account for the effect of the fish ladder on overall recovery. However, it may not always be possible to link the effectiveness of individual restoration projects with the overall recovery of injured resources or damaged services. A connection between effectiveness of a restoration activity and population recovery does not necessarily need to be a prerequisite to a successful and useful monitoring program.

In order for the information from the overall goals identified in number 2 above to be useful to decision makers and the public, it will be necessary for all monitoring programs to follow guidelines on standardized reporting units, data format, QA/QC, etc.

The process used to guide implementation and management of monitoring should also include frequent involvement of an independent, rotating pool of resource, monitoring and quantitative experts and statisticians. The same reviewers should not have responsibility for repeatedly reviewing a program. Review of monitoring efforts should include review of the monitoring protocol prior to implementation, review of restoration activities that have monitoring elements, ensuring that the objectives are addressed and the program is technically sound; and review of draft and final products resulting from both these efforts.

1.5 PLAN ORGANIZATION AND CONTENT

This plan is divided into nine sections, beginning with this section, Section ?, 1, Introduction. In the introduction we provide the background for the program, elements of the conceptual plan, and our design and approach in developing the plan. Additionally, in Section ?, 1 we provide summarized responses to the specific questions of our scope of work agreement. This is followed by a discussion on the value of monitoring recovery (Section 2), and definitions of recovery (Section 3). Section 4 covers the needs and objectives of the recovery monitoring program as well as the strategies to meet these. Next, in Section 5, the injured resources and damaged services are identified along with monitoring recovery endpoints for each. Several of the resources and services whose endpoints are perhaps less definitive than others, are discussed in this section. Section 5 also provides a mechanism for prioritizing monitoring activities, with the development and application of criteria. This section also provides information on the linkages between resources and services. Both can be further developed with the development of resource- and service-specific conceptual models, as recommended. Guidance on sampling design, with resource- and service-specific information, is provided in Section 6. Section 6 also provides information on other monitoring programs that may be useful to integrate and/or coordinate with, and programs to learn from. Management of the monitoring program is covered in Section 7, overall recommendations in Section 8, followed by the last section, Section 9, the references cited.

2. WHY MONITOR?

2.1 VALUE AND USES OF MONITORING

Why should the Trustee Council devote funds from the *Exxon Valdez* oil spill settlement to monitoring recovery? Given the many demands for these funds, this is an essential question to answer.

Monitoring will allow the Trustees to:

- Measure the success and rate of recovery of resources and services in the spill area.
- Determine the natural recovery rate and effectiveness of selected restoration projects.
- Facilitate resource and service restoration.
- Establish a starting point for future comparisons and improve on existing baseline information to aid in detection of and response to effects of future oil spills or other perturbations.
- Serve as a long-term damage assessment.
- Assure the public that recovery of resources and services is occurring.

Monitoring's greatest value to the Trustee Council is the public assurance and documentation that the injured resources and damaged services are recovering. The *Exxon Valdez* oil spill and subsequent spills world wide have produced concern and fear among many people. This fear includes the perceptions that resources will never recover; the Trustee Council is not capable of ensuring recovery of resources and services; restoration will occur whether recovery demands it or not; and that settlement funds are being used to support activities that will not yield results. Well-designed monitoring activities, that are designed and coordinated with one another, are a vital element as they will assure the public that recovery is occurring. Development of a conceptual monitoring plan is the first step, lending objectivity to the process of monitoring recovery.

The Trustee Council's responsibility for stewardship of the natural resources requires them to ensure that resources injured and services damaged by the oil spill are recovering. This can only be accomplished through monitoring. Monitoring must be sufficiently rigorous and scientifically defensible to provide confidence to the public and the scientific community that recovery is documented.

Part of the overall goal for monitoring is to identify previously undocumented injuries that may exist. Through recovery monitoring the Trustees will provide a vehicle that may detect such injuries.

Perhaps more important, the information gather through recovery monitoring and through monitoring indicator parameters will provide a baseline, a lack of which proved a significant and overwhelming detriment to determining the extent and magnitude of the spill effects. Establishing a baseline for the future, along with documenting recovery and the effectiveness of restoration activities, should be perhaps the highest priority of the Trustee Council in their stewardship of the natural resources and services.

Monitoring results can be used for various purposes. Results of recovery monitoring may be used to assist in determining whether or not oil and gas development should be allowed and where. Results may provide information on the effects from clean up activities versus oiled areas. They may also aid in understanding and evaluating population dynamics, for instance, when do clean up activities help, hurt or make no difference to a resource or service. Monitoring the assisted recovery, or the effectiveness of restoration activities, provides a cause-and-effect evaluation of how useful the restoration activity may be in other spill situations. Lastly, long-term monitoring can provide information useful to the various user groups. For instance, parameters for long-term monitoring should include those that pertain to risks (such as changes in dynamics), those compelling to the public, and those with broader implications to the ecosystem as a whole.

2.2 CONSTRAINTS ON MONITORING

The main constraint on monitoring is monetary (i.e. availability of funds). To monitor each of the injured resources and damaged services throughout the entire geographic area of the spill and throughout several generations would be cost prohibitive. This necessitates consideration of the actual costs and the public's perceptions of the costs of the project, and benefits associated with selecting resources and services to monitor with the hope that the information can be extrapolated to other resources and across geographic areas.

Other constraints on monitoring include the general lack of baseline information for some resources and services, and practical considerations such as logistics, seasonality, etc. The lack of baseline information can in some cases limit the ability to statistically compare changes and estimate variation. In some cases, control sites can be used in place of, or in addition to, pre-spill information. Practical considerations may preclude monitoring of some resources or services at particular sites and during particular times of the year.

Logistical constraints such as weather, tides, extensive geographic area, remoteness of some areas, etc., all put limits on the information gathered during a monitoring program. Additionally, scientific constraints, such as resources whose life cycles are not fully understood, whose habits are secretive and thus difficult to monitor, whose habitat is difficult to work in (e.g., underwater), and whose populations were on decline prior to the spill for either a known or unknown reason, all affect monitoring and the information obtained and how it's interpreted.

Even if recovery endpoints are identified for a particular resource or service, it may not be possible to monitor that resource or service due to an inability to quantify the endpoint, or an inability to monitor the resource or service. In other words, some endpoints may not be able to be monitored.

Another constraint is the effect a particular monitoring or restoration activity may have on another activity. This is relevant for program elements within this monitoring program, as well activities in other programs that may effect the activities and/or results of this program. This emphasizes the need for coordination of both inter- and intra-specific activities. Additionally, restoration activities designed and implemented to assist the recovery of one resource or service, may actually negatively impact the recovery of another resource or service. Similarly, monitoring activities that may not disrupt the resource being monitored, but may disrupt another species.

3. DEFINITIONS OF RECOVERY AND LONG-TERM MONITORING

3.1 RECOVERY

Recovery is a term that means something different to different people. Recovery of the various natural resources, and the services they support, following the *Exxon Valdez* oil spill will occur at variable rates for different resources and will likely vary geographically across the spill area and between populations. Thus various degrees of recovery will be present in different resources and services and at different locations in the future.

It is necessary to define the term recovery. For the purposes of the conceptual monitoring plan, the term recovery means a return to "normal" or estimated levels or limits of what current populations/conditions would be had the spill not occurred ("no-spill conditions"). Recovery of resources and services can occur through natural biotic and geomorphic processes (except archaeological resources) as well as through restoration or manipulation of existing conditions to facilitate recovery. Recovery of services may also include replacement or enhancement of affected resources and services, or elements thereof.

For specific resources and some services, recovery to predicted "no-spill" levels may not occur for many generations, if ever. For instance, the draft Restoration Plan identifies natural (unassisted) recovery of injured resources to range from four to 120 years, with "unknown" listed as time to recovery for six of the 18 resources listed. Examples from the draft plan include: archaeological resources cannot recover at all; black oystercatchers may recover in 15 to 30 years; recovery estimates for marbled murrelets are unknown. Additionally, it may take ten years to discern actual recovery from natural variation or background noise. Other factors (stresses), both natural and anthropogenic, influence resources, services, and ecosystems. Resources and services respond to multiple stimuli and the response to anthropogenic influences becomes superimposed over natural variability in a manner that could preclude generalizations from species to species, habitat to habitat, and service to service. Thus, a return to pre-spill or no-spill conditions may not be realistic or feasible. Recovery will most likely be the acceptance of some steady state of conditions, an equilibrium that takes into account natural variation, that may differ from those that existed before the spill.

Ideally, "complete recovery" of resources would include:

- Presence at the locations had no spill occurred, in the abundances had no spill occurred, with the population age-class structure had no spill occurred, the biomass had no spill occurred, the linkages with other resources/parameters (i.e., same prey items) had no spill occurred.

"Complete recovery" of services would include:

- Use of the damaged area by the original user groups, to the use levels had no spill occurred, with the attitudes had no spill occurred.

Enhancement of a resource or service may also occur through on-going restoration activities. Enhancement goes beyond recovery. For instance, establishment of a population beyond the estimated no-spill level, or number of users increased beyond the no-spill levels.

In order for monitoring results to be used as an effective management decision-making tool, it is necessary to establish monitoring and recovery endpoints for resources and services to be monitored (see Section 5.3). Toward this end it may be necessary to define an achievable or "acceptable" level of recovery that may be less than ideal or complete recovery.

Because baseline or pre-spill information was not available for many resources and services, clearly defining the original conditions for some resources or services may not be feasible. Thus, we must identify other criteria for evaluating recovery. Pragmatically, recovery will be evaluated by investigation of only a sample of the species, habitats, and services affected by the oil spill and over a limited geographic area. Thus, it may be necessary to identify key taxa and representative services to monitor that can adequately assess a spectrum of the injured resources and damaged services.

3.2 RECOVERY MONITORING

Recovery monitoring is both the monitoring of natural, unassisted recovery of injured resources and damaged services, and the monitoring of specific restoration activities designed to aid recovery of a resource or service (personal communications, RPWG 1993). The primary focus of this plan is on recovery monitoring (i.e., natural unassisted recovery and/or the effectiveness of restoration actions of injured resources and damaged services), and on determining when recovery has occurred. The overall goal of recovery monitoring is to monitor the rate of recovery. Elements of recovery monitoring may extend into long-term monitoring. Additionally, general parameters, such as climatic data, identification of stresses, etc., may be elements of both.

3.2.1 Natural Recovery Monitoring

Monitoring of natural recovery is the monitoring of resources and services whose recovery has not been knowingly assisted through anthropogenic manipulation. In some cases natural recovery will be indistinguishable from assisted recovery or assisted recovery will be an element of overall recovery. For instance, the effect of the installation of a fish ladder on a particular stream to assist in the recovery of sockeye salmon may be measurable within that particular stream system, but its effect on the sockeye population as a whole, throughout the spill area may not be distinguishable.

3.2.2 Restoration Monitoring

Several restoration activities that involve anthropogenic manipulation to assist in the recovery of resources and services have a monitoring component to determine if they are effective. This monitoring plan does not cover these existing monitoring programs; however,this program needs to be coordinated with any restoration monitoring efforts. Restoration monitoring covered by this plan will evaluate the effectiveness of specific restoration activities on aiding the recovery of selected resources and services. The decision on which restoration activities need to be monitored will be based on the Trustee Council's review of on-going and future proposed restoration studies. Those selected for potential monitoring can then be reviewed in light of the objectives and strategies described in Section 4.

Restoration activities and monitoring may act as anthropogenic stresses to the species they are meant to assist, as well as to other injured resources or services. In evaluating restoration activities to implement and/or to monitor, the effect on other resources and services should be considered.

3.3 LONG-TERM MONITORING

One goal of long-term monitoring is to provide information on existing spatial and temporal conditions, natural variation, and existing socioeconomic data such that changes due to future perturbations can be detected. The priority for collecting this type of information through long-term monitoring would depend, in part, on the perceived or actual need for that type of information. Collection of additional existing data should be guided, in part, by determining which types of data were lacking and would have been useful in determining the extent of injury or damage. Presumedly, this type of information would be useful in the future to evaluate the effects of a future perturbation. Another aspect of long-term monitoring utilizes indicator measurements to monitor the overall health of the ecosystem. Prioritizing long-term monitoring activities associated with evaluation the overall health of the ecosystem would, in part, be a function of identifying indicator and/or keystone species that would provide the greatest amount of information for the least cost and effort. These indicator measurements should detect change (e.g., sensitive physical, chemical, biological, and/or social, cultural and economic parameters) in which a change would be indicative of perturbation. In addition, long-term monitoring could also detect residual spill effects and provide ecological baseline information to assess the impacts of future disturbances. Long-term monitoring is defined here as monitoring that occurs over a five-year period, or longer, that collects data following long-term trends in the distribution and abundance of injured resources and the quality and quantity of services. In general, recovery monitoring elements all have the potential to become long-term monitoring elements, or indicators of ecosystem health and of future perturbations.

Data collection for recovery monitoring and long-term monitoring may overlap or be the same at times. Overall planning by the Trustee Council can take advantage of the overlap and give preference to those monitoring projects which fulfill multiple monitoring goals and objectives.

Additionally, as mentioned, in the description of recovery monitoring, other monitoring parameters such as climatic data, may be elements of monitoring, regardless of the monitoring type. Parameters for long-term monitoring should include those that pertain to risks, such as changes in dynamics, those compelling to the public, and those with broader implications to the ecosystem as a whole.

4. NEEDS, OBJECTIVES, AND STRATEGIES

Monitoring is essential to understanding if settlement activities have been successful at restoring, enhancing or replacing injured resources and damaged services. The overall goal or need (used interchangeably) of monitoring is stated in the draft Restoration Work Plan (1993). The overall goals are to develop a comprehensive and integrated monitoring program that will:

- Follow the progress of natural and assisted recovery
- Establish an ecological baseline from which future disturbances can be evaluated

These goals have been further broken down into specific needs, objectives, and strategies below. The objectives reflect the input from many individuals, including Restoration Team and Restoration Planning Work Group members, peer reviewers and principal investigators. The objectives, as stated, are comprehensive and need to be further refined in Phase 2 of the monitoring program, when the bounds (e.g., physical, financial) of the monitoring program are set.

The following list and prioritization or sequence of needs, objectives, and strategies of the conceptual monitoring plan reflects the general consensus derived from the interviews and the workshop. Section 4.1 outlines needs, objectives, and strategies that pertain to both types of monitoring (e.g., recovery *and* long-term), while sections 4.2 and 4.3 present needs, objectives, and strategies that are specific types of monitoring.

4.1 GENERAL MONITORING PLAN

1. Need

   Scientifically and publicly credible acceptable monitoring program.

   Summary of Need

   The monitoring program will be scientifically and publicly credible only if the individual projects are well thought out, planned, executed. Variability and uncertainty can be dealt with and minimized by the use of preliminary studies or historical data, reliable sampling, and analytical methodologies. The plans for the individual monitoring projects need to be subject to peer-review prior to project initiation and periodically throughout the project. All projects should also meet specified quality assurance and quality control (QA/QC) guidelines.

Objective

- Ensure a credible monitoring program, that if at all possible, limits the monitoring to testing hypotheses and sets limits on sample variability and account for natural variability for program elements. [Monitoring activities that cannot test hypotheses should explicitly state what they intend to accomplish and identify the problem and question(s) they intend to address.]

Strategies

- Specify monitoring requirements in the Request for Proposal (RFP), such as submittals must be formulated with testable hypotheses.

- Utilize a timely peer-review system to review proposals and reports for scientific credibility and merit, technical feasibility, including their ability to detect change, and how useful the data are to resource managers and the public.

- Review monitoring proposals and assess methods and reports to ensure that, whenever possible, testable hypotheses are stated and uncertainties (i.e., sample and natural variation) are addressed.

- Where needed, develop, or request development of, methods for monitoring.

- Develop a framework for QA/QC.

- Take public opinion and perception into account in developing the monitoring plan.

- Establish forums (e.g., scientific, community and agency participants) to evaluate effectiveness of monitoring studies.

- Establish a design and evaluation team of statisticians and modelers to provide a uniform, high level of expertise to those that will conduct the monitoring.

2. Need

- An accessible and/or integrated, well-designed database.

Summary of Need

Accessibility of the data is critical for the monitoring to be of any value to resource managers, scientists, and the public. In order to be an effective tool for decision-makers and investigators, a catalog of the monitoring data, as well as other spill related data, should be centrally located and accessible by the various user groups. A centralized cataloging system will allow for the past, ongoing, and future data collected using *Exxon Valdez* oil spill money to be accessed to maximize the information gained from the spill and to allow for comparisons between and within resources and services. Additionally, the database must be designed properly for easy retrieval of data useful to scientists, agencies and the public.

Objectives

- To have knowledge of and access to existing *Exxon Valdez* monitoring, damage assessment, and restoration data.
- To have knowledge of existing monitoring and resource management data that may be useful in understanding recovery of resources injured and services damaged by the oil spill.
- Ensure being able to access and retrieve monitoring data by the various user groups.

Strategies

- Identify and build an efficient structure with well-defined variables/fields, headers, linkages, selection tools, reporting forms, etc.
- Develop a well-designed centralized, computerized catalog or library of databases that should include, but not be limited to, contact name/agency, parameters measured, resource or service studied, and when possible, the summary statistics calculated.
- Code existing and future *Exxon Valdez* oil spill databases with a common link for location/site and resource or service so that information on resources or services is retrievable by a unique identifier, as is information on a location/site.
- Provide guidelines to principal investigators for standardizing components such as resource or location/site codes and reporting units, for ease in adding and retrieving data.
- Utilize a well-designed system that is user-friendly and provide step-by-step instructions on how to access and retrieve information from the catalog of databases.
- Determine the interface tools necessary.

- Design a flexible system to accommodate additional fields and respond to unforeseen needs as new information becomes available.

- Identify the potential needs of the user groups, including oil spill response teams, NRDA researchers, principal investigators, and public users.

- Identify an individual person to oversee the centralized catalog, including acquisition of databases and programming.

- Ensure the information is centrally located to facilitate its accessibility.

- Integrate the database with interpretive and analytical tools (i.e., routines/programs that allow retrieval of information in formats useful to users).

3. Need

- Information for long-term management of injured resources and damaged services.

Summary of Need

Monitoring results provide a tool for decision-makers to determine which resources and services are recovering on their own and whether or not the rate of recovery is acceptable, which may never recover, and which may recover with human assistance.

Objective

- Provide information useful to decision-makers.

Strategy

- Collect long-term data documenting recovery of injured resources and damaged services.

- Ensure accessibility of monitoring data to resource agency managers and other decision-makers, investigators, and the public.

- Develop models to evaluate the data in forms that are useful to various users.

4. Need

- Establish a link between project approval and funding for that project.

Summary of Need

A link between project approval and project funding needs to be established in order that a program designed to determine if recovery is occurring is not prevented from being implemented due to a funding shortage part way through the program. The project approval decision process needs to include steps for guaranteeing funding with feedback mechanisms that still allow for project review.

Multiple years of monitoring will be necessary in many cases to ensure that injured resources and damaged services have recovered. Recovery of several of the resources may not be detectable within a ten year period due to a variety of factors (e.g., time to reproductive maturity and fecundity). Due to this constraint, guarantee of a long-term funding source needs to be established prior to implementation of some monitoring programs.

Additionally, even for resources where recovery can be measured in less than 10 years, the programs will likely involve multiple year studies, and/or periodic monitoring. To ensure that funding will be available to complete studies requiring periodic monitoring over several years, it will be necessary to establish a link between project approval and funding that ensures a long-term funding mechanism. One funding link or method is to establish an endowment to fund activities after Exxon payments end.

Objective

- Fund multiple years of monitoring.

Strategy

- Establish an endowment to be used for multiple years of recovery and long-term monitoring after Exxon payments end (i.e., greater than 10 years).

5. Need

- Consistency and timeliness in data reporting.

Summary of Need

To maximize the usefulness and compatibility of the data obtained through monitoring, standardization of reporting requirements and ensuring the timely submittal of results is necessary.

The guidelines developed will not dictate what methods investigators must employ to study their resource or service, rather the more general aspects to follow, such as reporting data in metrics, utilizing one of five possible software packages as a database software, etc.

Objectives

- Provide proposal and reporting guidelines (covering components such as publishing requirements, standardization of units, use of convertible software, status reports, QA/QC requirements, ideas on statistical methods to employ, etc.).

- Establish a method for ensuring timely submittal of deliverables.

Strategies

- Require periodic one page progress reports and project end reports with date of deliverables dependent on the resource- and/or service-specific studies.

- Develop guidelines (covering components such as publishing requirements, standardizing units, convertible software, status reports, QA/QC requirements, ideas on statistical methods to employ, etc.) for principal investigators to follow.

- Develop recommendations for Request for Proposal and contract language that sets specifics for reporting and schedule commitments and penalties.

- Establish general reporting requirements for information potentially useful to a variety of programs, such as collection of climatic data.

6. Need

- Program design that provides a feedback mechanism and integration with other monitoring programs.

Summary of Need

Throughout the monitoring, feedback mechanisms will be important to ensure that monitoring is effective at determining if recovery is occurring at an adequate rate, and to ensure coordination/integration with existing monitoring programs and others as they come on line. These mechanisms should be instituted at the design phase of the monitoring to ensure they are accomplished and there is no duplication of effort.

Objective

- Establish a method for ensuring feedback/evaluation of the monitoring program, and for coordination/integration with other programs.

Strategies

- As a proposal requirement for monitoring elements require that the submitter identify existing programs to coordinate with and how they propose to accomplish this.

- As a proposal/contract requirement institute a feedback/evaluation process to ensure that the monitoring element is attaining its objectives.

7. Need

- Dissemination of information to the user groups.

Summary of Need

Although not necessarily a component of the monitoring program, for the recovery monitoring result to be useful, the results must be available to the users.

Objective

- Identify a mechanism for timely dissemination of information that is available and understandable by the various users.

Strategies

- Through the NRDA process and ongoing restoration activities including public comments, generate a list of the user groups and the type of information they need (e.g., summary information, data on specific resources and services, etc.).

- In the proposal/contract development, require that respondents agree to the submittal of summaries of their programs, reports and data at scheduled intervals and in a set format, attend forums to share information, identify data from other monitoring elements that would be useful to them (e.g., mussel contamination data may be useful to those studying sea otters).

4.2 RECOVERY MONITORING

Recovery monitoring covers both assisted and unassisted recovery of injured resources and damaged services. However, some monitoring needs are specific to restoration activities, thus the recovery monitoring needs have been divided below into those addressing both assisted and unassisted recovery, and those specific to assisted recovery (effectiveness of restoration activities).

4.2.1 Monitoring of Natural and Assisted Recovery

1. Need

- Prioritization of resources and services to monitor, and elements thereof.

Summary of Need

Given that monitoring funding resources are finite, a series of decisions will determine how comprehensive and integrated the monitoring program will be. The recovery or restoration of some important resources and services may not be able to be monitored due to the physical properties of the system, biological properties of organisms, or logistical constraints in the area.

Objective

- Develop a method for prioritizing resources and services, and the monitoring activities.

Strategies

- Develop selection criteria to prioritize resources and services to monitor.
- Utilizing teams of experts, and the consensus-building process, establish priorities for recovery monitoring of selected resources and services by evaluating how well injured resources and damaged services meet criteria.
- Evaluate prioritization of monitoring programs in light of public opinion/perception (Phases 1 and 2 of the monitoring program).
- Develop criteria to identify resource- and/or service-specific monitoring activities (e.g., the life stage, behavior attribute, or population dynamic) and sampling designs (including statistical review) that are likely to document the success or failure of recovery (Phases 1 and 2 of the monitoring program).
- Evaluate potential monitoring activities through utilization of population models.
- Obtain cost estimates for conducting specific monitoring activities (Phase 2 of the monitoring program).
- Evaluate prioritization of monitoring activities in light of their cost-effectiveness to ascertain the quantity and quality of information to be gained versus the costs to be incurred. As necessary, reprioritize monitoring activities accordingly (Phase 2 of the monitoring program).

- After determining a common benefit currency (e.g., time to endpoint) evaluate the cost-effectiveness of recovery monitoring options (Phase 2 of the monitoring program).

2. Need

- A mechanism to document recovery of injured resources and damaged services.

Summary of Need

To monitor recovery over a long time period (e.g., 10 years or greater), some monitoring projects should be designed in serially repeating phases. These projects could continue as long as deemed necessary to determine if recovery has occurred, providing satisfactory work was completed. Satisfactory work would be defined independently of the results obtained. Some of the resources near oil spills in cooler temperate climates show significant effects of the spills at least ten years after the event (Ballou, T., et al. 1989; Chan, G. L. 1977; Clark, R. C., et al. 1978; Conan, G. 1982; Cretney, W. J., et al. 1978; Dauvin, J-C 1987; Dauvin, J-C. and F. Gentil 1990; Elmgren, R., et al. 1983; Gulliksen, B. and J. P. Taasen 1982; Jacobs. R. 1980; Linden, O., et al. 1979; Notini, M. 1978; Teal, J. M. and R. W. Howarth 1984). Provisions should be made for selective projects to continue for many years. Long-term monitoring could also occur by monitoring at periodic intervals of several years duration.

The preliminary assessment of damages has already occurred and will be used as a basis for defining recovery monitoring projects. It should be recognized that additional unsampled and presently undiagnosed damage effects may be discovered, and they may need to be included in the monitoring plan at a later date. Numerous monitoring alternatives need to be examined for each project. These include, but are not necessarily limited to, timing of sampling, types of sampling, geographical area to be examined, specific parameters to sample, and logistical effort necessary to accomplish the project.

Additionally, the monitoring program should be flexible enough to alter and add projects as new data becomes available. Although monitoring of some resources will serve as indicators for a large number of other resources, those indicators may not necessarily be determinable prior to the initiation of the sampling program. Initially, many resources and service may need to be monitored in a given area, with the number of resources or services monitored being reduced as data are analyzed to allow a sharper focus on fewer resources and services.

Objectives

- Establish a monitoring program to document the recovery of resources and services.

- Design a flexible monitoring program to accommodate rededication of efforts as new information becomes available.

Strategies

- Based on input from resource experts and/or population dynamic specialists, establish what acceptable rates of recovery are for each resource and service, or elements thereof (Phase 1 and 2 of the monitoring program).

- Utilizing a team of statisticians, identify appropriate intervals (monitoring frequency) for determining recovery of a resource and service over time and space (Phase 2 of the monitoring program).

- Determine the influence of other perturbations (natural or anthropogenic) on recovery (e.g., winter kill, other die-offs, predation, human disturbance, climatic changes such as El Niño, commercial fishing pressures, etc.) (Phase 2 and 3 of the monitoring program).

- Utilize existing data for assessment of baseline conditions (pre-spill, and/or damage assessment and restoration control site data).

- Utilize existing data (from the spill and from other programs) for developing recovery monitoring methodologies.

- Implement a periodic review system that allows for rededication of efforts as new information becomes available.

- Involve scientific experts and resource and service specialists during development of the monitoring program (Phase 1, 2, and 3 of the monitoring program).

- Develop a monitoring scope that encompasses the strategies above.

3. Need

- Knowledge that recovery is occurring, and the rate of recovery (endpoints).

Summary of Need

In order for recovery monitoring to be an effective tool there must be measurable endpoints -- measures of the rate and acceptability of recovery for each monitored resource and service. There may be multiple endpoints for some resources and services.

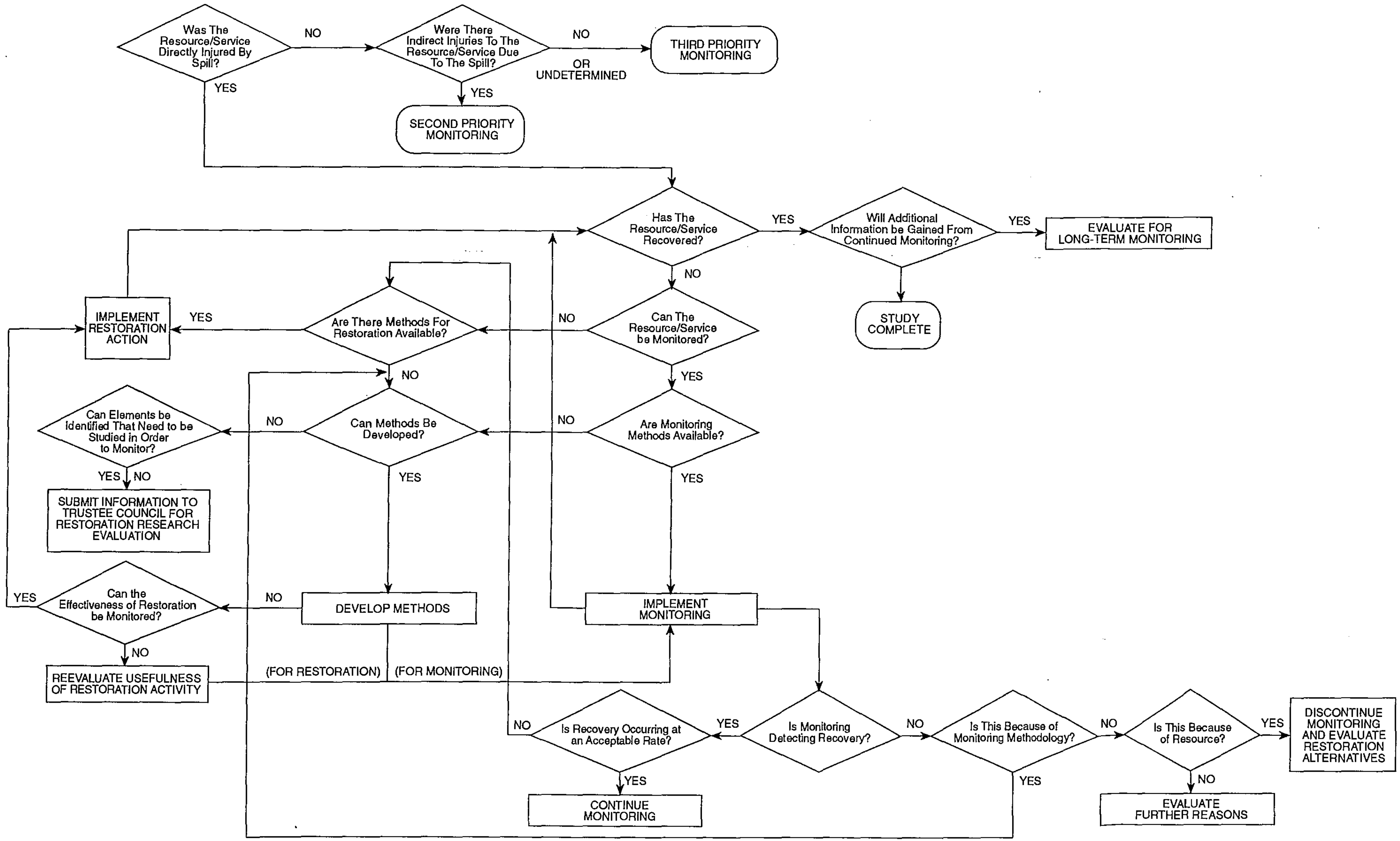


Figure 4.
Example of a Decision Tree Illustrating the Application of Selection Criteria to the Resources/Services

For any particular resource or service, the pre-spill, control condition, or perception and value will be defined as best as possible by multiple resource and service experts and/or existing data. The information that will be used to define the endpoint(s) should, whenever possible, include some quantitative measure of central tendency, such as a mean, median, or mode, and some indication of variance. For some resources or services, such quantitative measures may not be available or possible to define. In these cases, the information available will be used to describe the pre-spill condition, service, or resource, and this shall serve as the indication of the condition.

It may not be feasible to monitor recovery of some resources or services to a level comparable to the pre-spill conditions. Some resources or services may have been on the decline prior to the spill, and some may be so severely impacted that recovery is not possible within a reasonable time period (e.g., sea otter recovery estimates range from 15 to 50 years; common murre estimates range from 50 years to 120 years).

Still other resources or services may not become comparable to pre-spill conditions because of ancillary or unrelated changes resulting in an altered and non-comparable situation after the spill. For example, a resource may not achieve pre-spill abundance and distribution if other resources have increased to fill the carrying capacity based on a common food source. The carrying capacity having been reached, the injured resource would not be able to achieve pre-spill levels.

<u>Objectives</u>

- Define recovery endpoints for injured resources and damaged services.
- Evaluate whether resources and services are recovering at an acceptable recovery rate, as defined for each resource and service.

<u>Strategies</u>

- Based on input from resource and service experts, define recovery endpoints for each injured resource and damaged service, and periodically evaluate as data accumulate.
- With resource and service experts and statisticians, and economists establish what constitutes acceptable rates of recovery for each resource and service based on what is known about the resource and service.
- Compare the resource- or service-specific acceptable rate to the monitoring data obtained to reach a decision point: If rate of recovery is acceptable, evaluate need for continued or reduced monitoring frequency. If rate of recovery is unacceptable, evaluate restoration alternatives and/or research opportunities.

4. Need

- Establish linkages between resources and services in order to understand recovery.

Summary of Need

Although the tendency of monitoring is to focus on individual taxa, the *Exxon Valdez* oil spill had an impact on a large geographic area consisting of many different communities and trophic levels. By the very nature of the impacted areas, interactive and interdependent processes were disrupted, altered, or destroyed. Ecosystems are more than the sum of their parts, and the effects of perturbations such as the oil spill, can be experienced on the ecosystem level. The complexity of ecosystems, however, tends to render them difficult, if not impossible, to study as units. The study of recovery of such a relatively large association of altered communities (animals and human) could be not only difficult, but cost prohibitive. However, with the judicious choice of resources and services to be monitored, key components of the ecosystem's recovery can be addressed, and the recovery of the system as a whole may, in some instances, be inferred.

Objective

- Base the recovery monitoring plan on linkages between injured resources and damaged services that incorporates any knowledge of trophic levels interactions, and spatial and temporal variability.

Strategies

- Determine links and the types of interactions (e.g., positive, negative), wherever possible, between resources and services by evaluating available information.

- Develop resource- and service-specific conceptual models that include biological, physical, social, and cultural interactions and processes.

- Select resources and services for monitoring that are linked via trophic levels or that can be used to draw inferences about similar resources and services.

- Synthesize the information between resources and services and produce a concise summary.

4.2.2 Needs Specific to Monitoring the Effectiveness of Restoration Activities

It should be noted that monitoring the effectiveness of a specific restoration activity may also be an element of the restoration activity itself. Either way, monitoring the effectiveness of restoration activities should be integrated and/or coordinated with other types of recovery monitoring at the earliest stage possible, preferably at the design stage.

1. Need

- A mechanism to document effectiveness of select restoration activities at aiding the recovery of resources and services.

Summary of Need

To monitor the effectiveness of selected restoration activities, some restoration activities should incorporate a performance standard that specifically measures the ability of the activity to assist in the recovery of the resource or service (either at a local, population, or ecosystem level). These projects would be monitored as long as deemed necessary to determine if restoration has been effective in aiding recovery.

In addition, monitoring the effectiveness of select restoration activities could also identify where additional restoration activities may be appropriate, determine if delayed injury occurs, and determine if restoration activities for one resource or service are having a positive or negative effect on another resource or service.

Lastly, the evaluation of the effectiveness of restoration activities should be flexible enough to alter and add projects as new data become available. Although the monitoring of some resources and services may serve as indicators for a large number of other resources and services, those indicators may not necessarily be determinable prior to the initiation of the sampling program. Initially, many restoration projects may need to be monitored, with the number being reduced as data are analyzed allowing a sharper focus on fewer restoration programs.

Objectives

- Establish a monitoring program to document the effectiveness of restoration activities.

- Design a flexible monitoring program to accommodate rededication of efforts as new information becomes available.

Strategies

- Establish what is acceptable recovery for a resource or service during and/or after restoration implementation.

- Identify appropriate intervals (monitoring frequency) for determining effectiveness of restoration.

- Determine the influence of other perturbations (natural or anthropogenic) on restoration activity.

- Determine the influence (positive and/or negative) of the restoration activity on other resources and/or services.

- Utilize existing data for assessment of baseline conditions (pre-spill, and/or damage assessment and restoration control site data).

- Utilize existing data (from the oil spill and from other programs) for developing effectiveness of restoration monitoring methodologies.

- Implement a periodic review system that allows for rededication of efforts as new information becomes available.

- Involve scientific experts and resource and service specialists during development of the monitoring program.

- Develop a monitoring scope that encompasses the strategies above.

2. Need

- Knowledge that restoration activities are effective in aiding recovery, and the resulting rate of recovery is within the expected or estimated range.

Summary of Need

To determine whether restoration is effective in aiding recovery, there must be measurable endpoints — measures of the rate and expected or estimated rates of recovery for each monitored resource and service.

For any particular resource or service, the pre-spill or control condition will be defined as best possible by resource experts and/or existing data. The pre-spill conditions that will be used to define the endpoint(s) should, whenever possible, include some quantitative measure of central tendency, such as a mean, median, or mode, and some indication of variance. For some resources or services, such quantitative measures will not be available or possible to define. In these cases, the information available will be used to describe the pre-spill condition of the service or resource, and this shall serve as the indication of the condition.

It may not be feasible to monitor the effectiveness of restoration of some resources and services to a level comparable to the pre-spill conditions. Some resources or services may have been on the decline prior to the spill, and some may be so severely impacted that restoration does not aid recovery within a reasonable time period.

Objectives

- Define restoration endpoints for injured resources and damaged services.
- Evaluate whether resources and services are being restored at an acceptable recovery rate, as defined for each resource and service.

Strategies

- Based on input from resource and service experts, define restoration endpoints for each injured resource and damaged service.
- Establish what constitutes acceptable rates of recovery after restoration for each resource and service based on what is known about the resources and services.
- Compare the resource- or service-specific acceptable recovery rate to the monitoring data obtained to reach a decision point: If restoration results in an acceptable rate of recovery, evaluate the need for continued or reduced frequency monitoring. If restoration activities result in an unacceptable rate of recovery, evaluate continuing or selecting alternative restoration options.

4.3 LONG-TERM MONITORING

1. Need

- Identification of natural and anthropogenic stresses to aid in development and interpretation of monitoring elements.

Objective

- Identify potential stresses to resources and services.

Strategy

- Specify as a contract requirement (part of scope of work) that principal investigators include a reporting section discussing anthropogenic and natural stresses on the resources or services they are studying and how these might influence the results obtained.
- Develop resource- and service-specific conceptual models that include biological, physical, social, and cultural interactions and processes.

2. Need

- Information on natural, temporal, and spatial variation of indicators to allow identification of a catastrophic event (health of ecosystem) and reduce the impact of such perturbations.

Summary of Need

In order to detect change that is outside the range of natural variation it is necessary to establish the bounds of natural variation. Long-term monitoring is required to define these bounds. Once established, monitoring should then be able to detect changes that extend beyond the bounds of natural variation.

Objectives

- Develop a monitoring program to detect spatial and temporal changes in biological and/or physical parameters that fall outside the range of natural variability.

- Follow long-term trends to provide baseline information for future perturbations.

Strategies

- Review past and present trend monitoring programs to identify matrices/parameters useful in detecting environmental change.

- Review past and present damage assessment and restoration data to identify resources with population effects attributable to the oil spill.

- Evaluate which recovery monitoring programs should evolve into long-term monitoring programs.

- Select physical, chemical, and/or biological indicator matrices/parameters for monitoring temporal and spatial changes in environmental quality based on the following:

    - Parameters sensitive to perturbations (i.e., those that will show a change), and

    - Parameters that are well understood (i.e., a solid basic knowledge of natural variation, and/or thorough knowledge of life history).

- Evaluate ease (i.e., cost-effectiveness, ability to dove-tail with other studies, frequency of sampling required) of monitoring these parameters.

- Design and implement a program that encompasses the above strategies.

3. Need

- Identify and understand linkages between physical, biological, and/or chemical parameters, as well as social and cultural interactions.

Summary of Need

It is necessary to select indicator parameters for monitoring because it is not economically or logistically feasible to monitor all resources or services for long-term monitoring. Indicators should enable inferences to effects on other resources, service, or parameters, but first one must establish the linkages between the parameters.

Objective

- Enable inferences to be made concerning higher trophic level exposure/health.

Strategies

- Determine links, wherever possible, between parameters monitored by evaluating available data on interactions between physical, biological, and chemical features, including exposure mechanisms (the coupling of monitoring multiple trophic levels with process studies), as well as social and cultural interactions.

- Select parameters that are linked via trophic levels or that can be used to draw inferences about similar species or services.

- Evaluate selected parameters in relation to the geographic location and physical setting (e.g., enclosed embayment) to determine if they will be effective indicators.

5. RESOURCES AND SERVICES TO BE MONITORED

The settlement requires that use of restoration funds be linked to injured resources and damaged services resulting from the *Exxon Valdez* oil spill. The injuries summarized in the Restoration Framework (*Exxon Valdez* Oil Spill Trustees 1992) and recently completed damage assessments, along with input from the RPWG were used to prepare a list of injured resources and damaged services. Injured resources are further divided into those effected at the population level (direct effects), and those indirectly effected. The injured resources and damaged services are then to be prioritized for recovery monitoring.

5.1 RESOURCES

Resources injured at the population level are identified in the draft Restoration Plan and listed below:

- Mammals
  - Sea otters
  - Harbor seals
- Birds
  - Common murre
  - Marbled murrelet
  - Pigeon guillemot
  - Harlequin duck
  - Black oystercatcher
- Fish
  - Sockeye salmon
- Community Assemblages
  - Intertidal biota
  - Subtidal biota

Resources that were injured but did not appear to experience a population decline as a result of the spill include (also see draft Restoration Plan):

- Mammals
    - Killer whales
    - River otter

- Birds
    - Bald eagle

- Fish
    - Cutthroat trout
    - Dolly varden
    - Pink salmon
    - Pacific herring
    - Rockfish

Other injured resources include:

- Archeological sites and artifacts
- Designated wilderness areas

Other resources may have been injured either directly or indirectly as a result of the oil spill but were not studied during the NRDA process. The list of injured resources may change as monitoring results become available.

5.2 SERVICES

Damaged services identified as important to monitor include:

- Commercial fishing

- Commercial tourism
    - Tour ships
    - Day tours
    - Hunting and fishing charters

- Passive uses (also called aesthetic, wilderness, intrinsic or non-use value)

- Recreation
    - Sport fishing
    - Sport hunting
    - Motor boating

- Ocean kayaking
- Sailing
- Hiking and camping

- Subsistence.

5.3 RECOVERY ENDPOINTS

Recovery endpoints provide the measuring stick for evaluating whether or not recovery has occurred. The endpoints differ for each resource or service, and by definition of recovery. The vertical axis on Table 1 provides a list of resource and service endpoints developed through the workshop process described in Section ?, 1. The definition of recovery is based on the definition of conditions existing prior to the *Exxon Valdez* oil spill as discussed in Section 3, and further illustrated in Table 1. As shown, in some cases attainment of an endpoint is based on pre-spill conditions, in others, on control site monitoring, and still others, on the perception of the resource and service users (see Table 1). Controls can obviously be influenced by a variety of variables, particularly the mobility of the resource and the hydrodynamics of the area. A final category, long-term, indicates those endpoints that would be useful for long-term monitoring to detect future perturbations. Table 1 is only partially completed because the workshop participants did not include experts covering all of the injured resources and services presented along the horizontal axis. Nor did the workshop generally have more than one or two experts present for a particular resource or service. Therefore, Table 1 is considered draft. Further refinement of this table is suggested in Section 8, Recommendations, with the suggestion that at least three experts on each resource or service contribute to evaluation and selection of recovery endpoints.

It should be noted that not all recovery endpoints can be monitored. In other words, endpoints may define recovery for a resource or service but not be achievable. It will be important to identify these to the user groups, particularly the public, to aid in their understanding of why or why not a particular resource or service was monitored. It may also aid in development of monitoring methodologies to measure endpoints. Also, it should be noted, that once recovery of a resource or service is attained, that does not necessarily mean that monitoring of that resource or service should be discontinued. Continued monitoring may provide invaluable information on ecosystem health and/or on the effects of further perturbations, and thus be an element for long-term monitoring. The continuation of monitoring may also provide information on enhancement of resources and services beyond recovery. Continued monitoring beyond recovery will have to be evaluated on a case-by-case basis.

One potential method that could be used in conjunction with scientific monitoring results to determine when monitoring of resources and services should end, is to undertake a public opinion survey. When the public opinion survey indicates that people no longer are willing to pay for certain monitoring activities or restoration activities (i.e., an indication that people feel

Table 1. Matrix of recovery endpoints for injured resources and services.

	Killer Whale	Sea Otter	Harbor Seal	River Otter	Pigeon Guillemot	Black Oystercatcher	Common Murre	Marbled Murrelet	Harlequin Duck	Cutthroat Trout	Dolly Varden	Sockeye Salmon	Pacific Herring	Rockfish	Pink Salmon	Archeological Sites/Artifacts	Intertidal Communities	Subtidal Communities	Bald Eagle	Designated Wilderness Area	Commercial Fishing	Commercial Tourism	Passive Users	Subsistence	Recreation	Mussels	Forage Fish
Biological																											
- Population Size	●	●	●	●■	●■	■	●	●	●			●	●		●■		■	■	●							●■	
- Mean Population Size w/CVs	●	●■		■		■	●■	●	●			●			●■		■	■	●							■	
- Reproduction/Recruitment	●■	●■	■	■	■	■	●■		●■				●		●■		■	■	●								●■
- Growth (Individual, Physical) Rate		■					●		■	●	●				●		■									■	●■
- Physiological		■		■		■	■		▲■						●		■										
- Population Equilibrium (Population Growth)	●■	●■			■	■	●								●		■	■	●								●■
- Age Class Sex Structure	●■	●■					■	■	●			●	●		NA		■	■								■	●■
- Prespill Condition Adjusted for Change (Decline or Increase)		▲			●		●	●							●■		■	■	●								
- Mortality Rate	●■	●■					■		■						●■		■										
- Distribution (Density)	●	●		■		■	●	■	■			●	●		●		■	■	●								
- Behavior																											
- Habitat Usage	●			●■			●		●						●■		■	■									
- Community Structure (Diet Taxa Richness, Community Taxa Richness)	NA			●■			NA		NA						NA		■	■									
- Population Growth Rate							■		▲						●												
Physical/Chemical									■																		
- Eliminate Oil as Plausible Cause of Negative Effect				■		■			●■								●■	■	■	●◆	●◆	●◆	●◆	●◆	●◆	●■	
Services and Archeological Resources																											
- Reduction of Looting (Archeological)																●■▲					●	●	●	●	●		
- Usage Attained																											
- Hydrocarbon Concentration no Longer Effects Organic Components of Sites(Archeological)				●	■				●■							●■▲	●		■								
Achievement of Compensatory Action																											
- Quantity (Is it enough?)																				●◆	●◆	●◆	◆	●◆	●◆		
- Quality																				●◆	●◆	●◆	◆	●◆	●◆		
- Location																				●◆	●◆	●◆	◆	●◆	●◆		
- Perception	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	◆	●◆	●◆	●◆	◆	●◆	●◆		

● Prespill ■ Control[1] ▲ Trend ◆ Perception/Value

NA Not Applicable

[1] Control sites can be within or outside the spill area (e.g., historical). Sites within the spill area will be given precidence over sites outside the spill area. For use of sites outside the spill area there needs to be justification provided.

that the costs are exceeding the benefits), a decision needs to be made to stop the activity and redirect efforts.

There are some economic studies of damages to services completed by the Alaska Department of Law that may be helpful in determining and defining endpoints for services and some resources. It may also be useful to include economists in the process of determining endpoints for services to better understand the economic consequences of specific recovery monitoring and restoration activities. In addition, the Trustee Council may want to consider completing the NRDA economic study and comparing it with the Department of Law study to determine which damage assessment projects may be most useful to the recovery monitoring process.

If it is determined that a specific recovery endpoint cannot be measured, but the monitoring activity is necessary, make this issue explicit in the overall monitoring plan and the specific monitoring program.

5.3.1 Resources

The recovery monitoring endpoints for the resources and services identified at the workshop are presented in Table 1. The endpoints for two of the resources and all of the injured services, (those resources and services whose endpoints require information on the characteristics of the resource or service as well as social, cultural, and religious values, are further discussed below.

5.3.1.1 Archeological Resources

Archeological resources (i.e., archeological sites and artifacts) do not and cannot recover like other natural resources. Permanent damages to archeological sites and artifacts can occur if they are not restored. In general, the damages to archeological sites and artifacts occur through looting of sites and artifacts, erosion within and around sites as a result of clean up activities, and by oiling. Thus, "recovery" endpoints for archeological sites and artifacts are associated with the nature of the injury and tied directly to restoration activities.

Two endpoints were identified in the workshop: (1) reduction of looting of archeological sites and artifacts, and (2) hydrocarbon concentrations no longer affecting organic components of archeological sites. Both of these endpoints could be evaluated using pre-spill data, by establishing control stations and/or through long-term monitoring.

With respect to looting, expert opinions indicate that sites in the spill area that have not already been looted are likely to be looted in the future. Additional looting can occur because there is increased knowledge of location of sites as a result of clean up activities. In addition, graffiti on existing archeological sites and structures can elicit releaser cues that can promote additional looting. There is a need to remove existing graffiti and restore looter holes through direct physical restoration to prevent further damage.

5.3.1.2 Designated Wilderness Areas

The perception that wilderness areas within the oil spill area are no longer pristine resulted from the oil spill. The damage to this resource was a change in perception. The value of this perception for people is in knowing that the area is pristine even if the people never visit wilderness areas. Experts agree that regaining the original perception is not realistic. Thus, an objective recovery endpoint for this resource may not be definable. However, lack of a clearly defined endpoint should not preclude consideration of monitoring activities.

One activity suggested by experts that may sufficiently change peoples perception is to designate portions of Prince William Sound as wilderness areas. This activity may be beyond the immediate scope of the monitoring program and needs to be considered in the context of the entire restoration plan.

5.3.2 Services

Defining recovery endpoints for some services is difficult. A general recovery endpoint for services could be when there are no longer any reasonable casual links between the condition of the service and the oil spill.

5.3.2.1 Commercial Fishing

Several physical and biological factors along with fisheries management practices affect commercial fishing harvests. Determining damages to commercial fishery resources is affected by variations and fluctuations in the fishing industry and other practices (i.e, input from hatcheries). As a result, experts believe that designating one recovery endpoint for commercial fishing activities would be very difficult. Experts involved in the workshop identified two possible endpoints that could be related to commercial fishing. The first relates to eliminating oil as a possible cause of negative effects on commercial fisheries. This could be accomplished using pre-spill data and through an evaluation of the users perceptions and values associated with commercial fishing. The second endpoint relates to attaining levels of use similar to use levels before the oil spill. This endpoint could be evaluated using pre-spill data.

5.3.2.2 Commercial Tourism

There are several forms of commercial tourism that were damaged by the oil spill. These include tour ship cruises, day tours, and hunting and fishing charters. At the workshop, experts indicated that one possible endpoint that could be applied to all of these commercial tourist activities is for reservations and bookings with companies that provide these activities to return to pre-spill levels.

5.3.2.3 Passive Uses

Damages associated with passive uses of the environment are difficult to demonstrate and quantify. Recovery endpoints associated with passive uses of the environment need to be defined by characteristics of passive uses and perceptions and values that people place on the environments that provide opportunities for passive uses.

5.3.2.4 Subsistence

Damages to subsistence harvests are relatively well defined. There are concerns with contamination of resources among individuals and communities dependent on subsistence harvests as a livelihood. Two endpoints were identified at the workshop that could be considered in determining recovery of subsistence harvests. The first is to eliminate oil as a possible cause of negative effects on subsistence resources. This could be evaluated using pre-spill data or hydrocarbon data, as well as through an evaluation of perceptions (i.e., satisfaction with the type and level of subsistence activities) among subsistence hunters. The second endpoint could be to attain use levels of subsistence resources similar to the use levels before the oil spill. ADF&G surveys of subsistence use could provide baseline or pre-spill information.

5.3.2.5 Recreation

Recreational activities consist of sport fishing, sport hunting, motor boating, ocean kayaking, sailing, hiking, and camping. In general, two recovery endpoints for recreational activities were identified by experts at the workshop. The first is to eliminate oil as a possible cause of negative effect on the resources that support the recreational activities. This endpoint could be evaluated by using pre-spill information and surveys to evaluate peoples perceptions about the resources that support the recreational activities. The second endpoint is to have the level of use by any given recreational activity in specific areas return to the same or similar use levels before the oil spill. Pre-spill data for some recreational activities could be used to evaluate the second endpoint. For example, there is some anecdotal information on pre-spill use levels of ocean kayaking that could be compared to post-spill ocean kayaking activities. Recovery endpoints specific to each type of recreational service should be developed by service experts in Phase 2 of the monitoring program.

5.4 VALUE AND USE OF CRITERIA

The Trustee Council will be faced with deciding which resources and services to monitor and with choosing specific monitoring activities. How will this be done? Given the demands for settlement funds and the number of resources and services that could be monitored, it is important to develop a tool for evaluating the potential range of monitoring activities. A list of criteria have been developed to assist the Trustee Council in deciding which resources and services should be monitored and which studies of these resources and service will meet the goals of the monitoring plan (Section 5.5).

The criteria can be used as both a planning tool and a decision making tool. As a planning tool, the criteria can be used by the Trustee Council to:

- Determine which of the injured resources and damaged services to monitor.
- Develop specific requests for proposals for monitoring activities.
- Evaluate and rank proposals received in response to request for proposals to monitor specific resources and/or services.

The criteria could also be used by respondents to the request for proposal in preparing a monitoring proposal. Any proposed monitoring activity should consider each criterion in preparing a monitoring plan.

As a decision making tool, the criteria will be useful by the Trustee Council in deciding if a particular monitoring program is documenting recovery. The list of criteria should be used to evaluate the results of the monitoring activities (either on an interim basis or at the end of a monitoring element) to determine if recovery is occurring. If recovery is occurring or has occurred, the Trustee Council can make decisions to:

- Continue funding the program.
- Continue funding the program with reduced sampling effort and/or over a different time scale.
- To discontinue funding.

If recovery is not occurring, the Trustee Council can use the criteria as a guide to:

- Evaluate the need to invest in restoration alternatives for the resource or service.
- Evaluate the need to continue recovery monitoring but with a different focus.
- Decide if a feasibility study is necessary to determine why the resource or service is not recovering.

Socioeconomic concerns may also be an element the Trustee Council reviews. In part, the socioeconomic criteria or value of a monitoring action would be what society is willing to pay for the information gained. If the monitoring information can be linked to a substantial improvement in the probability of avoiding damages and injuries from another catastrophic event, the information may be highly valued. However, if the information assists resource managers in making small improvements in the population size of an already abundant population of a species, the information may not be as highly valued.

5.5 CRITERIA FOR SELECTING AND EVALUATING MONITORING ACTIVITIES

Criteria are proposed to assist the Trustee Council in prioritizing monitoring activities. The criteria can be applied to each resource or service. Formulation of the criteria was based on verbal and written input from the Restoration Planning Work Group and Restoration Team, peer reviewers, and principal investigators. The list of criteria presented herein was further refined during the workshop (Table 2). All participants in the workshop recognized that the criteria are not perfect, and will require further refinement during Phase 2 of the monitoring program. The criteria represent a tool that can be used by the Trustee Council to prioritize monitoring activities, plan monitoring activities and to make decisions on the effectiveness of the recovery monitoring efforts. The criteria are a series of statements related to the severity of injury or damage, capability of monitoring, importance of the resource or service, and other parameters (Table 2). An example of a decision tree developed from the criteria is illustrated in Figure 4. The question "can the resource or service be monitored?" can be further broken down, as illustrated in Figure 4, to allow prioritization and/or weighting of resources and services that can be monitored.

For each criterion, a ranking of high, medium or low can be applied. In order to ensure that the ranking is applied in a consistent manner, definitions for each of the ranks must be provided. Examples of possible definitions are provided below. In the example provided, a score of one indicates low, and three indicates high, unless otherwise noted.

For each of the criteria presented in Table 2, it may be useful to define what is meant by high, medium, and low, since the meaning for magnitude of injury and, for instance, socioeconomic importance, will most likely be different. Examples for these two criteria are provided below.

<u>Magnitude of Injury</u>

High: A high score, indicated by a three, indicates there has been a population level effect across more than one resource or service grouping, (e.g., colony, pod, archeological site), and across more than a one geographic area (e.g., the spill effected populations regardless of geographic area versus only the colony on Montague Island was effected).

Medium: A medium score, indicated by a two, indicates there has been a partial population or indirect effect to the resource or service, in that one colony or site was effected, but populations in other areas were either not effected or are recovering.

Low: A low score, indicated by a one, indicates that there has been an indirect effect or an unknown effect to the resource or service, or that a very limited population was effected and is already recovering.

Table 2. Criteria for evaluating resources and services for monitoring.

Primary

- Severity of Injury:
  - Magnitude
  - Is the Injury Continuing
  - Evidence of Recovery
  - Lack of Prespill Baseline
- Capability of Monitoring:
  - Testable Hypotheses
  - Restoration or Compensation Detectable, Quantifiable
  - Quality of Reference Data (Prespill or Control)
  - Logistics
  - Quality of Endpoint
  - Precision/Accuracy (Future Monitoring)
- Resource/Service Importance:
  - Socioeconomic
  - Cultural/Religious
  - Ecological

Secondary

- Contribution to Understanding Analogous R/S
- Limited Applicability to Fishing and Subsistence
- How Non-Destructive are Sampling Techniques
- Regulatory Restrictions Inhibit Monitoring
- How Well are Service Characteristics and Use Dynamics Understood
- Sources of Stress Known/Evaluated
- Ease of Integration/Coordination with Other Monitoring Programs
- Provide Data for the Evaluation of Future Perturbations
- R/S Monitoring Not Duplicated (at Necessary Precision/Accuracy) by Another Agency
- Restoration or Compensation is Benefit to Other Injured Resources or Services
- Achievement of Compensatory Action
- Quality
- Quantity
- Location
- Perception

Socioeconomic Importance

High: A high score, indicated by a three, means that the resource or service is very important socioeconomically by providing, for example, a livelihood or food source to a human population greater than 5,000.

Medium: A medium score, indicated by a two, means that the resource or service has some socioeconomic value but that it's value alone, either is of value to small numbers of people, or its value is limited unless grouped with other resources to form an overall high socioeconomic value. An example might be the value of Harlequin ducks to the tourism industry versus the value of seabirds and marine mammals combined.

Low: A low score, indicated by a one, means that the resource or service has little or no known socioeconomic value. In other words, for a resource, it is not a significant food or pelt source, nor a resource particularly sought after by the tourism industry.

5.5.1 Criteria for Evaluating Resources to Monitor

The criteria for evaluating resources and services to be monitored are divided into primary and secondary criteria (see Table 3). Agreement was reached on the division of criteria into primary and secondary categories during the workshop. Primary criteria are thought to be most important in evaluating which resources and services to monitor, secondary criteria provide additional important information to refining the selection. There are three primary criteria: (1) Severity of Injury, (2) Capability of Monitoring, and (3) Importance of the Resource or Service. Each of the primary criteria are broken into subcriteria, as listed in Table 2. The combined or mean rank of the subcriteria provide an overall rank of the primary criteria. There are seven secondary criteria (see Table 2). An example resulting from the workshop, of how the criteria can be applied to resources is presented in matrix table format in Table 3. Because services may not be adequately represented by the broad categories, such as recreation and commercial tourism, these have been further broken down and presented in a separate matrix table covering strictly damaged services (Table 4). The criteria presented on the matrix tables for resources and services are the same; however, some of the criteria apply solely to some resources or services.

To complete the ranking of the various resources and services, experts in each resource and service should be consulted. At least three experts on each resource or service should be asked to rank the resource or service based on the criteria and meaning of a high, medium and low ranking. For the purposes of discussion, we are presenting results from the workshop, during which participants were asked to go through the ranking process. However, there were not experts to cover all resources and services, nor were there generally multiple experts covering a single resource or service. The resulting ranking reflects the best

Table 3. Workshop example of application of criteria to the injured resources.

	KILLER WHALE	SEA OTTER	HARBOR SEAL	RIVER OTTER	PIGEON GUILLEMOT	BLACK OYSTERCATCHER	COMMON MURRE	MARBLED MURRELET	HARLEQUIN DUCK	CUTTHROAT TROUT	DOLLY VARDEN	SOCKEYE SALMON	PACIFIC HERRING	ROCKFISH	PINK SALMON	ARCHEOLOGICAL SITES/ARTIFACTS	DESIGNATED WILDERNESS AREA	INTERTIDAL COMMUNITIES	SUBTIDAL COMMUNITIES	BALD EAGLE	MUSSELS	FORAGE FISH
PRIMARY																						
Severity of Injury: (Mean Score of Subcategories)	1.8	2.7			1.3	1.8	2.7	2.2	2.8							2.5		2.7	2	1	2.7	1
- Magnitude	2.5	3			1	1.5	3	2.5	2.5							2		3	2	1	2	1
- Is the Injury Continuing	1	2			1	2	3	2	3							3		3	2	1	3	1
- Evidence of Recovery (3 = low, 1 = high)	2	3			2	2	2	2	3							N/A		2	2	1	3	
Capability of Monitoring (Mean)	2.8	2.3			2.2	2.9	2.3	1.8	1.8							2.5		2.7	2.6	2.7	2.7	2.7
- Testable Hypotheses	3	3			2	3	3	2	2							3		3	3	3	3	3
- Restoration Detectable, Quantifiable	3	2			2	3	3	2	2							3		3	3	3	3	3
- Quality of Reference Data (prespill or control)	3	2			1.5	2.5	2	2	1.5							1		2	2	2	2	2
- Logistics (1 = difficult, 3 = easy)	2	3			2.5	3	1	1	1.5							2		2	2	2	2	3
- Quality of Endpoint	3	2			3	3	2	2	2							3		3	3	3	3	2
- Precision/Accuracy (future monitoring)	3	2			2	3	3	2	2							3		3		3	3	3
Resource/Service Importance (Mean)	2.3	3			1.3	1.3	2	2	1.7							2		2.3	1.7	2.2	2.3	1.7
- Socioeconomic	3	3			1	1	3	3	2							1		3	1	2	3	1
- Cultural/Religious	3	3			1	1	1	1	1							3		1	1	2.5	1	1
- Ecological	1	3			2	2	2	2	2							2		3	3	2	3	3
SECONDARY																						
Contribution to Understanding Analogous Resource/Service	1	1			1	1	3	2	2							2		3	3	2	1	3
How Non-Destructive Are Sampling Techniques	3	3			3	2	3	3	3							3		3	3	3	2	3
Regulatory Restrictions Inhibit Monitoring (many restrictions = 1)	3	2			3	3	3	2	3							1		3	3	2	3	3
How Well is Life History Understood?	2	3			2	2	3	1	1							1 *		2	3	3	3	2
Sources of Stress Known/Evaluated	2	1			2	2	2	3	1							1		2	2	2	2	1
Ease of Integration/Coordination with Other Monitoring Programs	3	2			2	2	3	3	2							3		3	3	2	3	3
Provide Data for the Evaluation of Future Perturbations	3	3			3	3	3	3	3							3		3	3		3	3

* "Life History" of Arch. Resources is seen as now much we currently know about the resources in the oil spill area.

Table 4. Example matrix of injured resources for application of criteria.

	COMMERCIAL FISHING	COMMERCIAL TOURISM	Tour Ships	Day Tours	Hunting & Fishing Charters	PASSIVE USES	SUBSISTENCE	RECREATION	Ocean Kayaking	Sport Fishing	Sport Hunting	Motor Boating	Sailing	Hiking
Severity of Injury:														
Magnitude														
Is the Injury Continuing														
Evidence of Recovery														
Lack of Prespill Baseline														
Capability of Monitoring:														
Testable Hypotheses														
Restoration or Compensation Detectable, Quantifiable														
Quality of Reference Data (Prespill or Control)														
Logistics (i.e., Difficult, Easy)														
Quality of Endpoint														
Precision/Accuracy (Future Monitoring)														
Resource/Service Importance:														
Socioeconomic														
Cultural/Religious														
Ecological														
Contribution to Understanding Analogous R/S														
Limited Applicability to Fishing & Subsistence														
How Non-Destructive are Sampling Techniques														
Regulatory Restrictions Inhibit Monitoring														
How Well are Service Characteristices & Use Dynamics Understood														
Sources of Stress Known/Evaluated														
Ease of Integration/Coordination With Other Monitoring Programs														
Provide Data for the Evaluation of Future Perturbations														
R/S Monitoring Not Duplicated (at Necessary Precision/Accuracy) by Another Agency														
Restoration or Compensation is Benefit to Other Injured Resources or Services														
Achievement of Compensatory Action														
Quality														
Quantity														
Location														
Perception														

judgement of a variety of experts covering any of the resources or services they felt comfortable addressing (see Table 3).

After calculating the ranking of the primary criteria the information can be evaluated by comparing the results obtained for the various resources and services, or within taxa groups, such as fish, birds and mammals. We do not recommend summing the score as this may result in a bias, whereby a resource with a high total score may actually be the result of several low ranking items totaled. A presentation tool that may assist in interpretation is the use of three-dimensional graphs, such as those presented in Figures 5 and 6, illustrating values presented in Table 4. Another mechanism for ranking may be to convert the results to percentage responses, thus avoiding summation of the results.

Results of application of the secondary criteria can be utilized to clarify and/or supplement the results of the application of the primary criteria.

Figures 5 and 6 can be interpreted as those resources and services that resulted in the highest mean for each axis (importance on the vertical axis, capability of monitoring on the horizontal axis, and severity of injury on the axis providing depth) is given the highest priority for monitoring. This translates to those resources and services that are closest to the back and upper most point of the diagram. For instance, in Figure 5, the prioritization into grouping of resources and services to monitor follows (Note: not all injured resources and damaged services are reflected here because not all were ranked at the workshop):

First Priority:

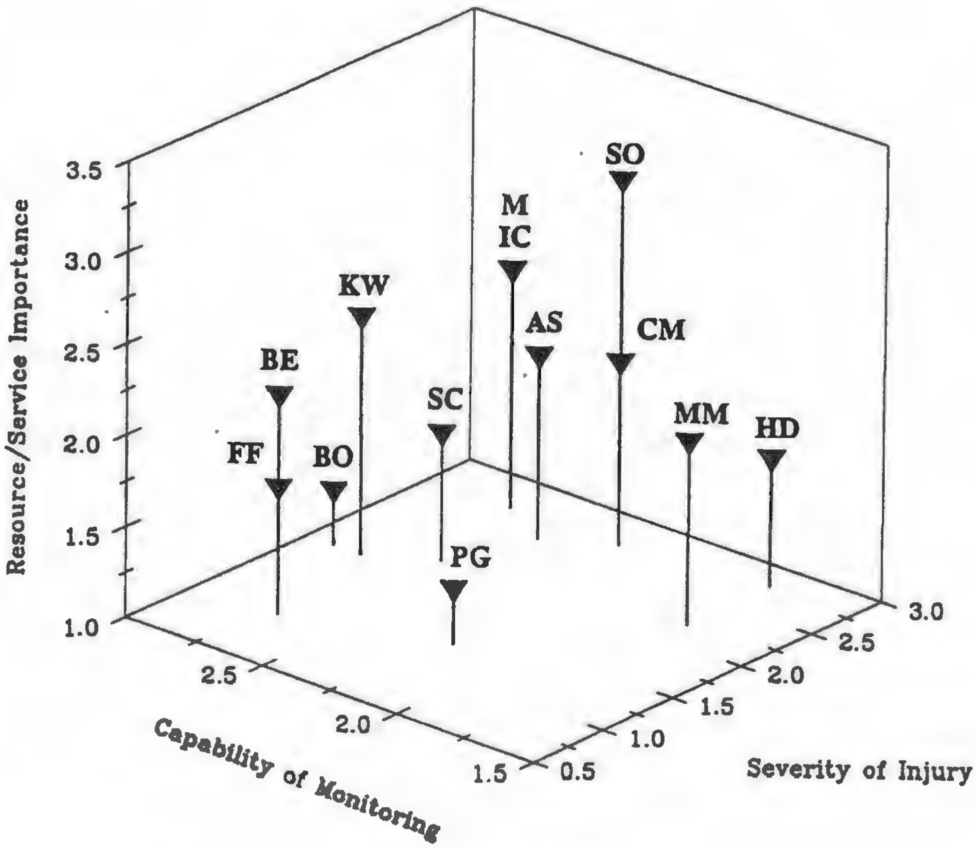
- Mussels and intertidal community
- Sea otter
- Archeological sites/artifacts
- Common murre

Second Priority:

- Harlequin duck
- Subtidal community
- Killer whale

Third Priority:

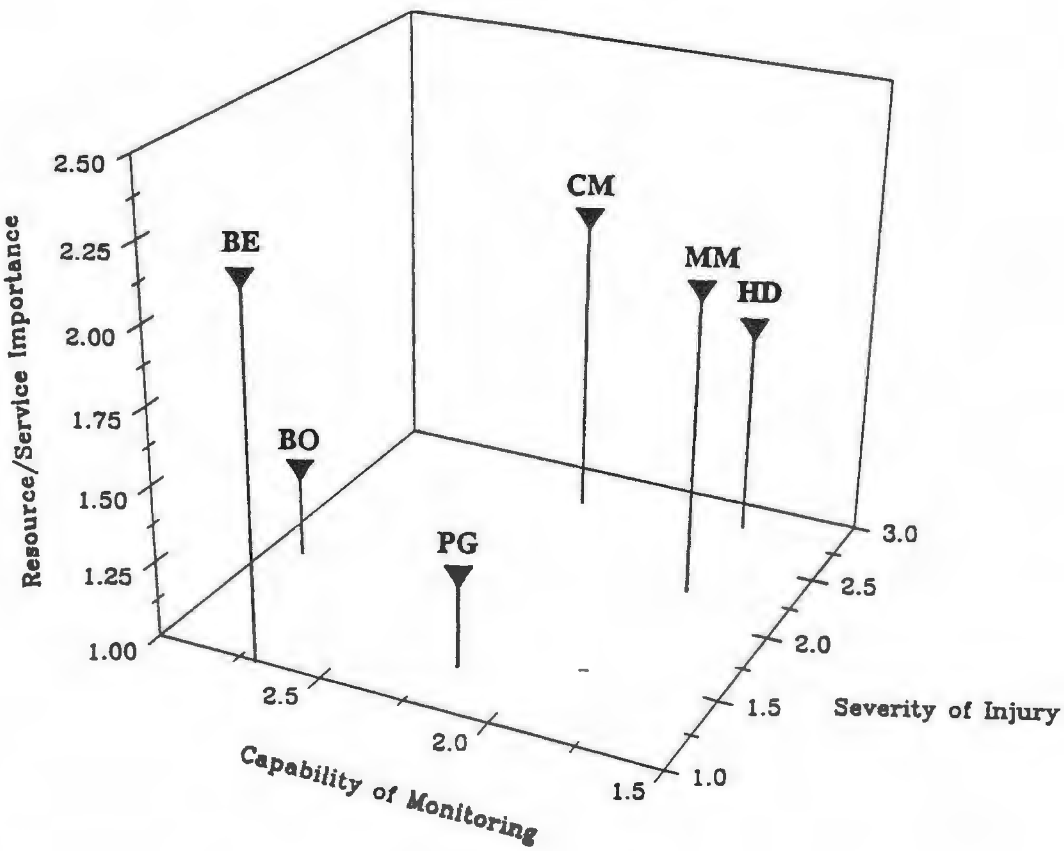
- Black oystercatcher
- Bald eagle
- Forage fish
- Marbled murrelet
- Pigeon guillemot



Key to codes:

AS=Archeological Site/Artifacts
BE=Bald Eagle
BO=Black Oystercatcher
CM=Common Murre
FF=Forage Fish
HD=Harlequin Duck
IC=Intertidal Communities
KW=Killer Whale
M=Mussels
MM=Marbled Murrelet
PG=Pigeon Guillemot
SC=Subtidal Communities
SO=Sea Otter

Figure 5.
Three dimensional graph illustrating results of application of primary criteria to the injured resources.



Key to codes:
- BE = Bald Eagle
- BO = Black Oystercatcher
- CM = Common Murre
- HD = Harlequin Duck
- MM = Marbled Murrelet
- PG = Pigeon Guillemot

Figure 6.
Three dimensional graph illustrating results of application of primary criteria to the injured bird species.

If a decision is made that in order to ensure that a broad range of resources are covered, and thus no single special interest is over-represented, the resources are divided into taxa groups, such as birds, mammals, fish and tidal communities. Reviewing the resources by taxa group may result in the example presented in Figure 6 for birds. In this example the prioritization of monitoring bird species follows:

First Priority:	Common murre
Second Priority:	Harlequin duck and marbled murrelet
Third Priority:	Black oystercatcher, bald eagle and pigeon guillemot

As mentioned earlier, Figures 5 and 6 present only examples of a prioritization process that could be used. These were developed at the workshop where there were a limited number of experts available, and no guidelines other than high, medium and low, provided for ranking resources and services. It may prove useful to the interpretation of the three-dimensional graphs to also specify the weight of a given axis. For instance, perhaps the severity of injury is the overall governing factor, which will allow for selection of resources and services based on similar results at other levels.

Referring back to the secondary criteria will also be useful in decision making. For instance, using Figure 6 (strictly the injured birds) and Table 2, it can be seen that the secondary criteria supported the conclusion that common murres should be a primary focus of monitoring. However, the secondary criteria do not aid in prioritizing the second and third level priorities. At this level, it may again, be necessary to attribute weighting to the criteria.

Another tool for use in evaluating resources and services to monitor is to review the linkages between resources and services. These are illustrated in Table 5, and will be further developed with the conceptual models during Phase 2.

Once criteria have been applied to the injured resources and services, they can be applied to the specific monitoring elements of the resources and services. Figure 7 is a basic example of the decision tree applied at this level.

For example, once a prioritized list of resources and services to be monitored is developed, requests for monitoring proposals can be developed for those resources and services with a higher priority for monitoring. The Trustee Council may receive five proposals to monitor the recovery of one highly-ranked resource or service. The criteria would then be applied a second time to each of the proposals for that resource or service. The proposals would be ranked and categorized into three categories (high, medium, and low). After all proposals for each of the top-ranked resource or service are ranked, all of the highest ranked proposals would be evaluated to determine any overlap between studies, identify opportunities for coordination between studies, and to determine if there are linkages between the different proposed studies that will assist in understanding recovery through trophic linkages.

Table 5. Matrix table of linkages between resources and services.

	Killer Whale	Sea Otter	Harbor Seal	River Otter	Pigeon Guillemot	Black Oystercatcher	Common Murre	Marbled Murrelet	Harlequin Duck	Cutthroat Trout	Dolly Varden	Sockeye Salmon	Pacific Herring	Rockfish	Pink Salmon	Archeological Sites/Artifacts	Intertidal Communities	Subtidal Communities	Bald Eagle	Designated Wilderness Area	Commercial Fishing	Commercial Tourism	Passive Uses	Subsistence	Recreation	Mussels	Forage Fish
Killer Whale		●	●									●	●		●			●			●	●	●		●		●
Sea Otter	●																●	●			●	●	●	●	●	●	●
Harbor Seal	●											●	●	●	●		●	●			●		●	●	●		●
River Otter														●			●	●					●		●	●	●
Pigeon Guillemot													●				●	●					●		●		●
Black Oystercatcher																	●	● Indirect					●		●	●	
Common Murre																		●	●		●	●	●		●		●
Marbled Murrelet																				●	●		●		●		●
Harlequin Duck									●			● (Roe)			● (Roe)		●	●	●	●			●	●	●	●	
Cutthroat Trout				●															●						●		●
Dolly Varden				●								●	●		●			●	●						●		●
Sockeye Salmon	●		●					●	● (Roe)		●		●				●	●	●		●			●	●		●
Pacific Herring	●		●		●						●	●		●			●	●			●			●	●		● (Larvae)
Rockfish			●	●									●					●			●				●		●
Pink Salmon	●		●						● (Roe)		●						●	●	●		●			●	●		●
Archeological Sites/Artifacts																						●	●		●		
Intertidal Habitat		●	●	●	●	●			●			●	●		●										●	●	●
Subtidal Habitat	●	●	●	●	●	● Indirect	●		●		●	●	●	●	●						●			●	●	●	●
Bald Eagle						●	●	●	●	●	●	●			●					●			●		●		●
Designated Wilderness Areas								●	●										●			●	●		●		
Commercial Fishing	●	●	●				●	●		●	●	●	●	●	●	● (Theft)		●						●	●		●
Commercial Tourism	●	●					●									●				●			●		●		
Passive Uses	●	●	●	●	●	●	●	●	●							●	●		●	●		●			●		
Subsistence		●	●						●			●	●		●			●							●	●	●
Recreation	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●	●
Mussels		●		●		●			●								●	●						●	●		
Forage Fish	●	●	●	●	●		●	●		●	●	●	● (Larvae)	●	●		●	●	●		●			●	●		

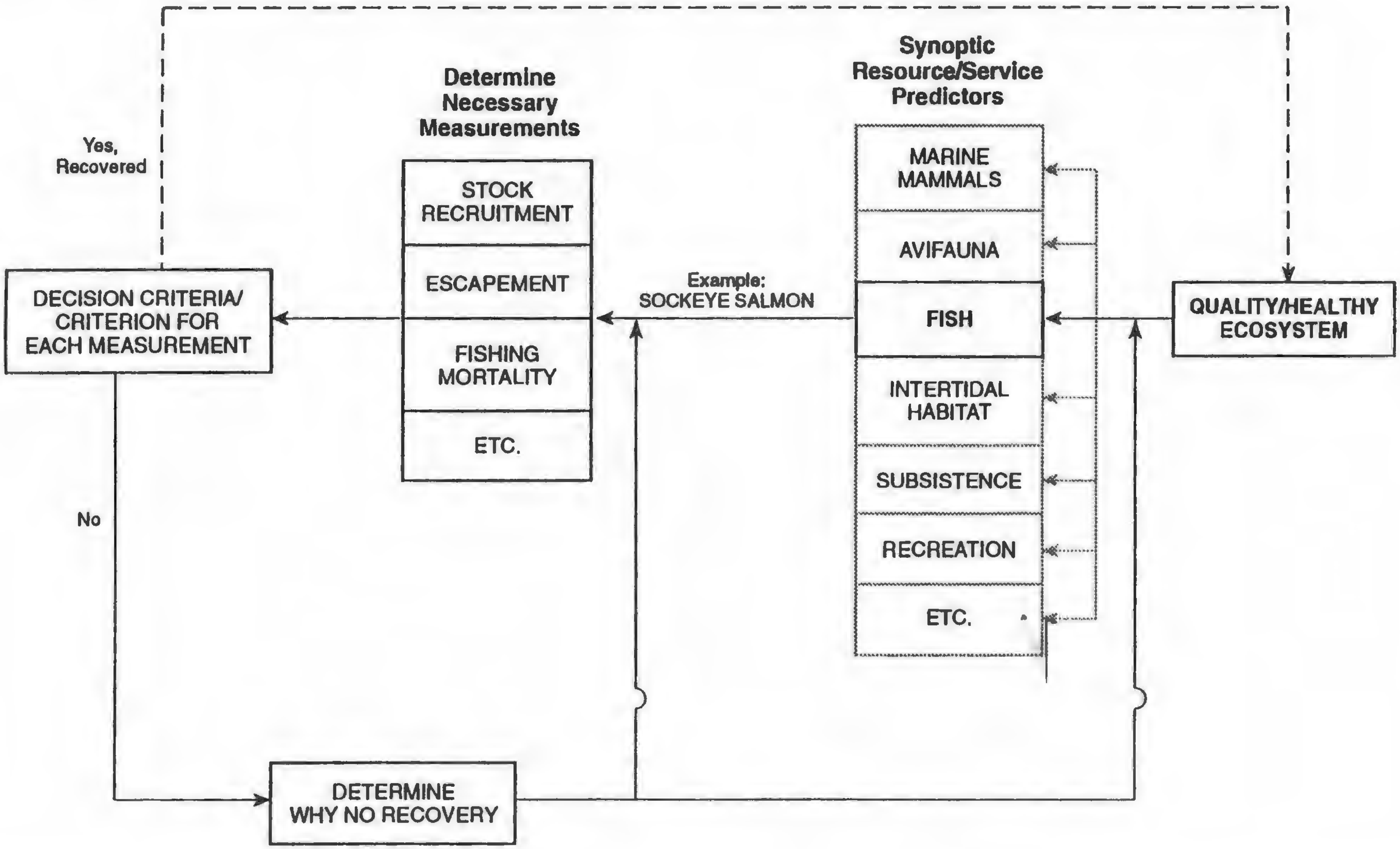
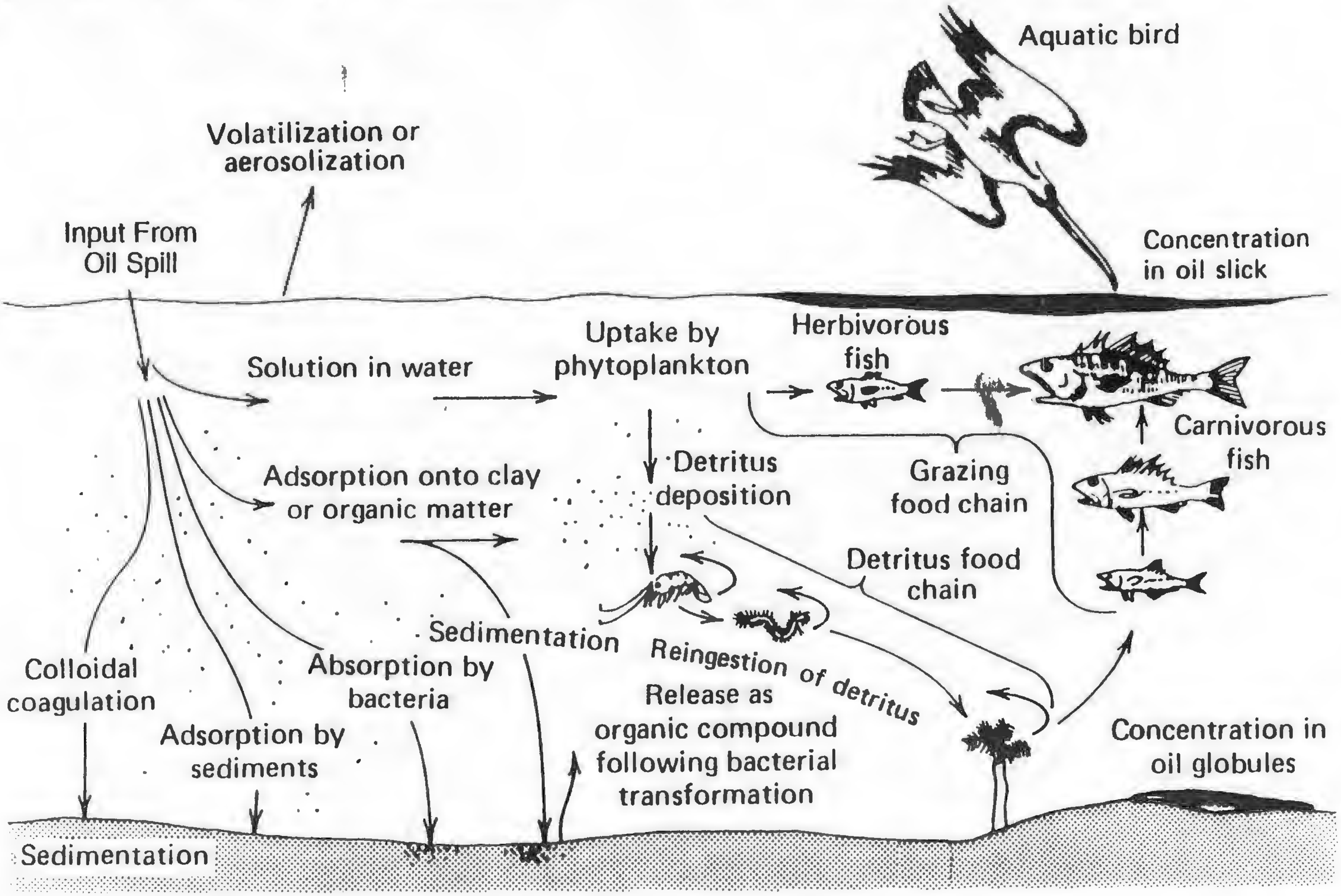


Figure 7.
Example of the Application of Selection Criteria to a Specific Resource Monitoring Element

Resources and services that do not receive a high rank during the first application of the criteria will not necessarily be eliminated from consideration in the future. Similarly, proposed studies that do not fall within the highest ranking category will not automatically be eliminated from consideration for funding.

5.6 DEVELOPMENT OF CONCEPTUAL MODELS

Conceptual models help define cause-and-effect relationships, aid in the development of hypotheses to test, and in understanding the interactions between biological, physical, and chemical interactions (Figure 8). As part of Phase 2 of the monitoring program, it is strongly recommended that conceptual models be developed for each resource and service. Figure 8 provides a generic example of a conceptual model. This would need to be further developed to address a specific resource or service and its interactions. Development of conceptual models will be assisted with the information provided in Tables 1 and 5, recovery endpoints and linkages between resources and services, respectively. The development of conceptual models can be completed by the contractor(s) for Phase 2, or as a requirement of a request for proposal and subsequent contract.



Source: Ecology, Impact Assessment, and Environmental Planning.

Figure 8.
Example of a Generic Conceptual Model for a Resource

6. GUIDANCE ON SAMPLING DESIGN

The design of monitoring elements should take into account methodologies developed to date. For instance, the methodologies employed in the NRDA and restoration science studies may be applicable to the monitoring program. In particular, the sampling stations, parameters measured, and units of measurements in these programs should be reviewed to optimize the information gained -- continuing the collection of data for comparative reasons. This same strategy should be used in evaluating monitoring programs other than this one, in order that the programs may be coordinated and/or integrated if the goals and objectives coincide.

Of course, the weaknesses in programs should be reviewed as well as the strengths. The NRDA, restoration science studies and other monitoring programs will surely provide lessons to be learned as well, such as information on an appropriate control, frequency of sampling, etc. Thus, just as the programs are reviewed for compatibility and continuation, they should be reviewed in order to strengthen as well as complement.

Below we provide general guidance on sampling design and statistical analyses. Section 6.1 covers the general principles to consider in the design of a program. This section is followed by guidance specific to resources (Section 6.2), with guidance given on the following taxa groups: avifauna and mammals, fish, intertidal and subtidal communities, and archeological resources. Guidance on monitoring services begins with Section 6.3, and covers each of the damaged services. Additionally, Section 6.3 covers potential programs with which to integrate and/or coordinate.

6.1 GENERAL GUIDANCE ON SAMPLING AND STATISTICAL ANALYSES

Statistics is playing an increasingly important role in environmental monitoring and research. This has been prompted by a need for valid and repeatable evaluations of elements of environmental (physical, biological, cultural) systems with known levels of confidence and uncertainty. Statistical sample design and analytical techniques can be employed to obtain rigorous descriptions of environmental conditions.

The purpose of this section is to summarize the statistical issues relating to any monitoring program designed to evaluate recovery from the *Exxon Valdez* oil spill. It is important that all monitoring programs use comparable techniques to design the programs, collect and analyze the data, and interpret the results.

Green's Ten Principles (Green 1979) outline considerations to the design of a defensible program.

1. Be able to state concisely to someone else what question you are asking. Your results will be as coherent and as comprehensible as your initial conception of the problem.

2. Take replicate samples within each combination of time, location, and any other controlled variable. Differences among can only be demonstrated by comparison to differences within.

3. Take an equal number of randomly allocated replicate samples for each combination of controlled variables. Putting samples in "representative" or "typical" places is *not* random sampling.

4. To test whether a condition has an effect, collect samples both where the condition is present and where the condition is absent, but all else is the same [may not be possible in the field]. An effect can only be demonstrated by comparison with a control [or a time series].

5. Carry out some preliminary sampling to provide a basis for evaluation of sampling design and statistical analysis options. Those who skip this step because they do not have enough time usually end up losing time.

6. Verify that your sampling device or method is sampling the population you think you are sampling, and with equal and adequate efficiency over the entire range of sampling conditions to be encountered. Variation in efficiency of sampling from area to area biases among-area comparisons.

7. If the area to be sampled has a large-scale environmental pattern, break the area up into relatively homogeneous subareas and allocate samples to each in proportion to the size of the subarea. If it is an estimate of total abundance over the entire area that is desired, make the allocation proportional to the number of organisms in the subarea.

8. Verify that your sample unit size is appropriate to the sizes, densities, and spatial distributions of the organisms you are sampling. Then estimate the number of replicate samples required to obtain the precision you want.

9. Test your data to determine whether the error variation is homogeneous, normally distributed, and independent of the mean. If it is not, as will be the case for most field data, then (a) appropriately transform the data, (b) use a distribution-free (nonparametric) procedure, (c) use an appropriate sequential sampling design, or (d) test against simulated H_0 data.

10. Having chosen the best statistical methods to test your hypothesis, stick with the result. An unexpected or undesired result is not a valid reason for rejecting the method and hunting for a "better" one.

Although evaluating testable hypotheses is a goal for monitoring, this may not always be possible. In such cases, the methods used should be established and thoroughly documented methods.

The first step in designing a specific monitoring programs is to define the purpose of the program (e.g., to determine if the population of bald eagles is recovering in Prince William Sound after the *Exxon Valdez* oil spill). This purpose will be used to develop the rest of the monitoring program, including what specific element(s) will be monitored to meet the purpose. Statistical theory and methods, in addition to knowledge of the characteristics of and influences on the resource or service to be evaluated, are used to guide the development and execution of the following components of a monitoring program:

- Formulation of testable hypotheses
- Statistical sample design issues
- What to sample
- Where to sample
- How to sample
- When to sample
- Statistical analyses
- Interpretation of results

These elements are addressed in more detail in the following subsections. Following Section 6.1, statistical considerations are discussed in terms of resource categories.

6.1.1 Formulation of Test Hypotheses

Based on the purpose defined for a specific monitoring program, a statement (the null hypothesis, H_0) is formulated that addresses the purpose in simple, concrete terms. This null hypothesis identifies the state of the element that is to be tested (e.g., if the purpose of a monitoring program is to determine if the population of bald eagles is recovering in Prince William Sound after the *Exxon Valdez* oil spill and the recovery monitoring endpoint is population size with an expected increase of two percent per year, the null hypothesis could be stated as "There is no statistically significant change in the number of bald eagles residing in Prince William Sound during the breeding season between 1993 and 1994.") Testing this null hypothesis via statistical methods will be the objective of the sampling and analysis process.

It is possible that the data will indicate that the null hypothesis is not likely true. An alternative statement (alternative hypothesis, H_A) is formulated that defines a different state of the resource or service. Should a statistical test indicate that the null hypothesis is false, the data can be evaluated in terms of the alternative hypothesis. For example, the alternative hypothesis for the null hypothesis in the previous paragraph could be stated as "There is a statistically significant increase in the number of bald eagles residing in Prince William Sound during the breeding season from 1993 to 1994".

When testing a hypothesis, as shown below, two types of error exist. Type I error (α), the probability of rejecting the null hypothesis when it is actually true, is commonly set at 5%. Type I error is also known as the significance level of a test. Type II error (β) is the probability of accepting the null hypothesis when it is actually false. Decreasing one of these two errors will increase the other.

		H_0 Accept	H_0 Reject
H_0	True	Correct Decision (1-α)	Type I Error (α)
	False	Type II Error (β)	Correct Decision (1-β)

6.1.2 Statistical Sample Design Issues

In order to develop an optimal sampling design for a monitoring program that tests the specified null hypothesis, some statistical issues must be addressed, including the significance level (α), power level (β), sources and magnitudes of variation, and minimum detectable change (MDC).

As noted in the previous section, two types of error are present in hypothesis testing, Type I (α) and Type II (β). These errors need to be balanced, since decreasing one increases the other. The only way to reduce one error level without increasing the other is to improve the sampling design, (e.g., increasing sample size). A sampling design must adequately and realistically address both types of error.

In environmental monitoring and sampling, many sources of variation exist in addition to those commonly addressed in experimental designs (e.g., within-sample, between-sample, analytical, and random). These additional sources also exist at very different magnitudes and dimensions. An optimal sampling design must also address temporal variation, spatial variation, and natural system variations. Replication is one sampling technique that can be used to quantify many of these sources of variation. Data from previous studies may also be useful in evaluating potential sources of variation.

When testing a hypothesis, a level of change exists below which the null hypothesis is not rejected. This minimum detectable difference (MDC) of a statistical test is affected by several other test parameters:

- Inherent variation (natural variation, within- and among-sample variation, and analytical variation)

- Sample size (n)
- Significance level (α)
- Power (1-β)
- Temporal and spatial autocorrelation

The minimum detectable difference should be small enough to meet the needs of the monitoring program but not so small as to require a prohibitively large sample size or reduce the power of the test below an acceptable level. Due to the amount of variability usually found in environmental data and limited sampling budgets, a balance between the MDC, sample size, significance level, and power has to be reached.

The following function can be used to study the balance between these quantities or evaluate the level of power associated with statistical tests under consideration for a single hypothesis.

$$(Z_{\frac{\alpha}{2}} + Z_{\beta}) * s * \sqrt{2\frac{(1-R)}{n}} * \sqrt{1+\rho(n-1)}$$

Where: $Z_{\alpha/2}$ and Z_{β} = The normal "Z" values for various levels of α and β

s = A quantity that estimates the inherent variation (commonly a standard deviation)

r = The temporal autocorrelation, or a quantity that estimates it

n = The sample size

ρ = The spatial autocorrelation, or a quantity that estimates it

A preliminary sampling effort should be made, if possible, to evaluate the design, evaluate the sampling and analytical procedures, and identify and quantify sources of variation. If the sampling design does not require extensive modifications, the preliminary data could be included in the monitoring program analyses. Another approach to address the variation issues would be to over-sample and use extensive replication the first two to three years to ensure adequate sample sizes and obtain estimates of variation components and then reduce the sample scope for the remainder of the monitoring program.

Any sampling design that is developed for a monitoring program should be flexible. It is quite possible that changes will have to be made after the first or second year to address inadequacies in the design or possibly budget constraints, especially if the amounts and primary sources of variation cannot be adequately assessed during the design phase.

6.1.3 What to Sample

After defining the purpose of a monitoring program, the resource or service to be sampled is decided so that the null and alternative hypotheses can be formulated. The state of the resource or service that is being tested should be relevant to the defined purpose of the monitoring program. There will probably be more than one characteristic of the resource or service that can be measured to test the hypotheses (e.g., the total number of bald eagles or the number of juvenile bald eagles). There should be little or no difficulty collecting data on the resource or service, and the data should have the capacity to evaluate the null hypothesis via statistical testing.

Choosing a characteristic of a resource or service ~~that~~ has been sampled in previous studies or monitoring programs may be advantageous. Data from these previous sampling efforts can be used to extrapolate properties associated with the measured characteristic prior to designing the sampling effort. Variability components could be estimated from the previous data to determine adequate sample size. Sampling and analysis problems encountered in prior work could be avoided or accounted for in the current study.

6.1.4 Where to Sample

Where to sample encompasses two issues, the study area and actual sample locations. The study area should encompass the entire area of interest for the monitoring program, and the sample locations will be sited within this study area. Previous studies can provide insight into appropriate methods and possible pitfalls.

The process of choosing the actual sample locations has implications to the statistical tests and their interpretation. How the sample locations are chosen influences the relationship between sample locations, variability estimates, and the inference basis for the statistical tests. Conventional statistical analysis methods were developed for data collected as random samples. Random, or probability, samples are considered independent and representative of the population from which they are sampled, and estimates of parameters such as means and variances computed from such samples are unbiased for those populations. By removing the randomness from the sample locations, as in judgment sampling, bias can influence the parameter estimates and restrict the interpretation of statistical tests.

There are three main sampling approaches that generate random samples: random, stratified random, and systematic random. In the random approach, samples are randomly located within the entire study area. In the stratified random approach, if the population of the resource or service under study is known or suspected to be unevenly distributed within the study area, homogeneous subgroups can be formed within the study area and random samples taken from each subgroup. The systematic random approach makes use of a two-dimensional grid that is randomly placed in the study area, and the sample locations are taken as either the intersections of the grid lines or at the same location in each grid area (e.g., the center). Other sampling schemes that produce random samples have been or can be developed from these basic approaches.

6.1.5 How to Sample

The methods used to collect data in the various monitoring programs should follow standardized protocols. These standardized protocols will ensure the data is consistent, accurate, and comparable between sampling events, monitoring years, and monitoring studies. Standardized protocols are also important to ensure that the data that will be analyzed is of acceptable quality for statistical analysis.

It is likely that many of the standardized protocols exist as a result of previous environmental studies. These will simply need to be assembled into a cohesive set. However, others may need to be developed from scratch, but previous research may provide useful insights into possible methods and difficulties.

Sampling methods should be documented in detail, specifying the exact steps to be taken from locating sample sites to shipping the collected samples to the analytical laboratory.

- Locating sample sites
- Collection of field observations (e.g., temperature)
- Collection of sample(s)
- Preparation of field spikes, duplicates
- Preserving, packaging, labeling of samples
- Storage, transportation of samples
- Documentation of samples (e.g., chain-of-custody forms)

Laboratory analytical methods should also be documented in detail; however, some of these will incorporate state and/or federal protocols.

- Receipt of samples from the field
- Preparation of samples
- Preparation of laboratory spikes, blanks
- Analytical procedures
- Reporting formats, including units and qualifiers
- Documentation of samples (e.g., chain-of-custody forms)

Compositing samples can be used to reduce the costs of analyzing large numbers of samples, increase the amount of sample material available for analysis, and reduce the between-sample variability caused by heterogeneous sample material. However, the consequences of compositing include the loss of ability to estimate between-sample variability, and this would need to be addressed in the sampling and analysis design.

6.1.6 When to Sample

Deciding when to sample is influenced by many factors. Natural factors, such as weather conditions and time of the year, influence the variability of the data that is collected. Organizational factors, such as sampling and analytical costs, can limit the amount and frequency of sample collection.

Historical data, if available, and knowledge of resource or service characteristics can be very useful in determining the best time(s) for sampling. Sampling times should be consistent from year to year (e.g., such as counting bald eagles present during the breeding season only). The severity of the weather in the oil spill area and constraints of monitoring (e.g., seasonal migrations of certain species) may also restrict sampling times.

Additionally, biological factors, such as life stage, behavior patterns, abundance and distribution of prey, etc., should also be factored into the decision on when to sample.

6.1.7 Statistical Analyses

There are many statistical analysis methods available for evaluating environmental conditions. The primary focus of monitoring programs is to detect change over time, and several standard analysis methods can be used to this end, such as analysis of variance (ANOVA), trend analysis, and time series analysis. Regression, correlation analysis, and other multivariate techniques can be used to evaluate hypothesized relationships between different variables measured in the monitoring program. Unless special circumstances require, standard analysis methods such as these should be used for the sake of clarity, comparability, and repeatability.

The analysis method used to test the null hypothesis is chosen prior to sampling, and it should be appropriate and rigorous for the stated null hypothesis and sampling methods used. Typically, the significance level (α) for a hypothesis test is set at 5%. For the chosen analysis method, the assumptions associated with the method must be addressed, since violations of an assumption can compromise the validity of and confidence in the analysis results.

Since spatial variability will most likely influence all monitoring data collected to some degree, statistical spatial analysis techniques may need to be considered. Any spatial methods used should be thoroughly researched and carefully applied. Geographic information systems (GIS) may be useful in such analyses.

A specific statistical method discussed at the workshop is a method that evaluates the year-by-year change in oiled sites relative to that at paired reference sites. There are several difficult analysis issues addressed by this method. This approach can be used for sites monitored on a yearly basis, or on a less often basis (e.g. every three years). By pairing sites, some large-scale variation can be accommodated inasmuch as the paired sites are proportionately affected. This method does not require any pre-spill data, which often does not exist.

While this method for evaluating the level of change between paired oiled and reference sites is straightforward and may be useful for any monitoring program, it does not replace the need for a sampling and analysis plan that is optimally designed to meet a specific monitoring goal.

This method is expressed as follows:

Design: n pairs of sites (oiled, reference) are monitored annually.

Data: In year t, $(T_{i,t}, C_{i,t})$ with $i = 1, 2, ..., n$

Incremental Relative Change:

$$\left\{\frac{\bar{T}_{t+1}}{\bar{T}_t} \Big/ \frac{\bar{C}_{t+1}}{\bar{C}_t}\right\} \begin{cases} > 1 \text{ suggests "recovery continues"} \\ < 1 \text{ suggests "damage continues"} \end{cases}$$

or, on the logarithmic scale, let

$$\bar{x}_t = \frac{1}{n}\sum_{i=1}^{n}\{(logT_{i,t+1} - logT_{i,t}) - (logC_{i,t+1} - logC_{i,t})\} = \frac{1}{n}\sum_{i=1}^{n} x_i t$$

so that

$$\sqrt{n}\,\frac{\bar{x}_t}{s_{x_t}} \begin{Bmatrix} > t_{1-\frac{\alpha}{2},n-1} \\ < t_{\frac{\alpha}{2},n-1} \end{Bmatrix} \rightarrow \begin{Bmatrix} \text{"recovery continues"} \\ \text{"damage continues"} \end{Bmatrix} \text{ else "don't know"}$$

Recovery: The sequence $\bar{x}_1, \bar{x}_2, \bar{x}_3$ ... converges to zero (in expectation)

Remarks:

- Serial correlation does not damage the individual t-tests, but does complicate testing for convergence.
- Pairing of oiled and reference sites is desirable, but not essential to this approach.
- The same approach would apply if sites were monitored, say, every third year.

- Annual "bay-wide" fluctuations are accommodated by this approach, to the extent that all sites are proportionately affected.

- Within-pair spatial correlation enhances power; between-pair distances should be great enough to make r=0.

- This method for data analysis does not use pre-spill data.

6.1.8 Interpretation of Results

Interpretation of results of analysis must take into consideration the tested hypotheses, sampling methods, and analysis methods and associated assumptions. The conclusions reached and interpretations made must be supported by the data and take into consideration the sampling methods used and the assumptions and restrictions associated with the analysis methods.

When drawing conclusions regarding environmental data, caution must be used. While a significant change may have been detected, can it be attributed to a recovery process or is it a result of some natural event (e.g., a decrease in predator population)? Because so many factors are not measured, conclusions regarding relationships between elements should be viewed as associations and not necessarily cause-and-effect relationships. Establishing cause-and-effect relationships in the environment requires controlling all factors not measured.

6.2 GENERAL GUIDANCE ON SAMPLING RESOURCES

Below are general guidelines for monitoring specific resource categories. Four general categories are specified: (1) avifauna and mammals, (2) fish,(3) intertidal and subtidal communities, and (4) archaeological resources.

6.2.1 Avifauna and Mammals

Throughout this general discussion of what, where, and how to measure, birds have been used to illustrate the point.

6.2.1.1 What to Measure

Some preliminary work is needed in order to determine what to monitor/measure. A primary consideration is to determine what questions need to be asked with regard to bird or mammal populations in the oil spill area. Development of these questions should rely heavily on previous work. If possible, new monitoring programs could be designed to be compatible with previous data collection such that meaningful comparisons are possible. When the appropriate questions (clear and unambiguous) have been asked, a monitoring program must be developed that will provide answers to these questions.

The monitoring program must be specific to the questions. For example, general questions on avian ecology may require censusing of large numbers of bird species and communities. For questions related to a specific food resource or foraging technique, a single representative species may be studied. For questions on reproductive success, surveys of breeding colonies and fledgling rates would be important. Toxicological questions requiring specimen analysis will have additional constraints on the selection of study species. Some species, such as bald eagles or marbled murrelets, may be important for study because of their status or public interest.

Further guidance on what to measure is available through use of the criteria presented in Section 5. The list of criteria should be reviewed to verify that they are appropriate for selection of suitable species. The criteria should also be relevant to the questions being asked. Applying these criteria to the list of species injured should lead to the selection of appropriate study species.

6.2.1.2 Where to Measure

Deciding where to measure a particular resource also depends on the questions that have been asked. If the objective is a long-term comparison with an existing pre-spill data set, it would be important to monitor in the same location and in the same manner as the previous work. If the question is how current conditions in the spill area compare with undisturbed areas, there would obviously be a set of parallel locations that differ only in their exposure to oil.

Where to measure would also depend on the species selected for study. For instance, some bird species can be studied most effectively in their breeding areas, while others are more easily studied in foraging areas. For some seabird species that nest far from the marine environment, such as marbled murrelet or harlequin duck, parallel studies might be needed in nesting and foraging or wintering areas.

When species and general locations have been selected, specific sampling sites should be chosen with a randomized procedure to ensure the statistical validity of the results. Likewise, an adequate number of replicate sites should be sampled to accurately estimate the means and variances of the variables to be studied. Simple random sampling could be appropriate for species that are known to be distributed homogeneously throughout the study area. For example, simple random sampling might be used for bald eagle nesting territories along the coastline or for area-wide censusing of foraging seabirds. Stratified random sampling would be more suitable for species with heterogeneous overall distributions that can be separated into homogeneous subgroups. For example, many seabirds nest in large colonies, each of which could be considered a homogeneous subgroup of the whole population. Specific colonies could be selected based on prior information, but sampling within a colony should be randomized.

6.2.1.3 How to Measure

Many measurement techniques are available, and the selection again depends on the questions being asked and the species being studied. If the objective is to duplicate previous data collection, identical techniques should be used, if possible.

Aerial surveys may be appropriate for broad-scale censusing, for instance of seabird concentrations, but these would not be suitable for species that are small, difficult to identify, or with few, widely scattered populations. Aerial surveys could provide very accurate estimates of bald eagle breeding territories and nesting success. For some species, different sampling methods would be necessary for different aspects of their life cycle. For example, boat surveys are needed to census foraging marbled murrelets, but ground surveys are necessary to locate possible nesting areas. Populations of some seabird colonies might be accurately determined from boat surveys, but in other colonies, sampling from the ground might provide better results.

Demographic studies of populations would require trapping and marking many individuals with distinct color bands. For studies on home ranges, it might be necessary to use radio transmitters on individuals. In small home ranges, a portable radio receiver on the ground would be most useful. In large home ranges or for long distance movements, it might be necessary to use helicopter or airplane-mounted receivers.

For physiological studies or for toxicological analyses, it would also be necessary to capture individuals. Some tests might be possible with samples collected in the field from animals that could be released. For example, blood and urine samples can be quickly and easily collected with no harm to a bird. Other tests, such as trace element analysis of organ tissues or electron microscopy of subcellular structures, would require the sacrifice of some individuals. Special review and approval should be required for studies of this sort.

6.2.1.4 When to Measure

Timing of sampling will often follow directly from the primary questions being asked and the species selected for sampling. Surveys of seabird breeding colonies would obviously have to be conducted during the breeding season. However, the breeding seasons are unlikely to overlap completely for all possible species of interest. Thus, prior information must be used to choose the optimum time for sampling.

Multiple sampling periods during a breeding season may be appropriate for some species. For example, a preliminary aerial survey could locate active bald eagle nests, and a survey later in the breeding season could determine the success rate of active nests. For other species, separate surveys in separate locations may be necessary for breeding and wintering populations. For example, harlequin ducks will breed on interior rivers and spend the winter in the near-shore marine environment.

The sampling season would also be important for some toxicological or physiological studies. Studies on hormonal changes related to breeding would have to be conducted over a time period spanning the breeding season. A study on trace element concentration in fat deposits would require sampling when body fat would be at a maximum.

6.2.1.5 Data Organization

As is the case with other aspects of a sampling design, the data organization is also dependent on the questions being asked. Data organization will also depend on the analytical procedures to be used, and the analysis should be considered in the initial phases of sampling design. Data organization must be carefully planned to be compatible with prior data sets and with the overall restoration database. However, it is possible that trying to make a data set conform to a general database might make it very difficult to use those data to answer the specific questions of a specific study. In those cases it would be most efficient to organize the data in a manner that will provide the most usable results for the specific study.

In general, it is anticipated that data would be organized in a matrix format. The simplest format would be a two-dimensional matrix with columns representing independent and dependent variables and rows representing individual measurements. Typical independent variables would be time, location or condition descriptors, and environmental factors. The dependent variables could be any factor that could be measured and would contribute to answering the initial questions. A three-dimensional matrix might be appropriate if similar, simultaneous studies are to be conducted on several species. Computational techniques allow the use of multi-dimensional data matrices, if that degree of complexity is appropriate to the initial questions.

6.2.1.6 How to Analyze

Data analysis is a key element of a sampling program, and it is essential to consider analysis when designing field data collection procedures. The analytical procedure must be focussed toward answering the basic questions of the study, and the data collection must be appropriate for the analytical procedures. If the proper data were not collected initially, the best analytical techniques will never produce meaningful results.

The analysis should consider the nature of the data in determining the appropriate statistical methods. Nominal variables are purely qualitative and cannot be assigned numerical values, and thus they may only be suitable for signs-based nonparametric or categorical statistical methods. Ordinal, or ranked, variables can be assigned numerical values, but the differences among ranks are not necessarily proportional. These variables must be analyzed with non-parametric statistics. Many ecological measurements will be discrete variables, which can have only integer values. Examples of discrete variables would be the number of bald eagle nests on an island or the number of seabirds on a transect. Discrete variables may be treated with parametric statistics, provided they satisfy the assumptions of those methods or can be transformed via a monotonic mathematical function to do so; categorical test methods may also be appropriate. Other environmental measurements are continuous, with no limits on possible values. Many physiological or toxicological measurements, such as metabolic rates or chemical concentrations, are continuous and can be analyzed with parametric statistics.

Depending on the statistical methods used for analysis, appropriate numbers of replicate samples must be taken to obtain accurate estimates of means and variances. If it is not possible to collect enough replicates, it may be necessary to use non-parametric statistics. However, without sufficient replication, statistical power may be compromised. Data available from previous studies can be used to evaluate variability and determine an acceptable level of replication.

A number of basic assumptions must be satisfied in order to use parametric statistics. It is essential that samples be taken at random. Errors in measurements must be independent and normally distributed, and variances of samples must be equal. Most parametric statistics assume normally distributed data; however, they are usually quite robust in the presence of non-normal. Unequal sample sizes and non-independent samples cause the most trouble. Statistical tests of these assumptions should be made to ensure the validity of the results. If the data do not fit a normal distribution, they may be normalized by a transformation. For example, most data on species populations are not normally distributed, but they can be normalized with a log transformation. If a suitable transformation for the data cannot be found, non-parametric statistics should again be used.

The particular statistical procedure to be used will obviously depend on the nature of the questions being studied. If the objective is to make comparisons between areas that were directly affected by the oil spill with other areas that were not affected, then a *t*-test or analysis of variance might be appropriate. To demonstrate functional relationships simple or multiple regression is possible. Trend analysis could be particularly useful in demonstrating recovery of a resource over time.

6.2.1.7 How to Interpret

If the objectives of the monitoring program were clearly stated, specific questions were addressed, and a comprehensive monitoring procedure was implemented, then interpretation should be straightforward. The results of the analyses should directly answer the questions that were asked, and reasonable conclusions should be drawn from the results.

Interpretation of the results of any particular study must be firmly based on reliable data that have been analyzed by statistically valid and relevant procedures. Any interpretation is only as rigorous as the weakest element in the entire data collection and analysis sequence. Care should be taken to avoid extrapolating to any interpretation beyond that which is justified by the available evidence and analysis.

6.2.2 Fisheries

6.2.2.1 What To Measure

The focus of fish sampling should be on those species that were injured by the oil spill. Specifically:

- Sockeye salmon
- Cutthroat trout
- Dolly varden
- Pink salmon
- Pacific herring
- Rockfish

These species provide opportunities to monitor ecological and biological variables as well as services provided by the resources. Monitoring activities on these species allow investigations at differing levels of the food chain. Further, the species differ in migratory behavior, life span, and exposure to the original oil spill. These differences allow investigations of recovery over a wide range of parameters from genetic integrity of populations to abundance.

Long-lived species, such as rockfish, may still show signs of spill-related impacts at either the population level (e.g., altered age structure) or individual level (e.g., physiological affects) and may be particularly valuable in assessing recovery.

With the exception of rockfish, the target species migrate through many habitats during their life history. The questions posed in the monitoring program must be carefully tailored to the species life history and niche in the aquatic community.

Each of these species provide direct services to humans and monitoring of these services may be appropriate to assess recovery and/or identify harvest management actions that may be desirable to speed recovery.

6.2.2.2 Where to Measure

The questions posed will guide the selection of sampling areas. However, due to the widespread distribution of the target species, it is important to focus the sampling efforts carefully to be certain of the level of exposure experienced by the population.

6.2.2.3 How to Measure

The primary determinant of the sampling methodology will be whether monitoring of biological parameters or services, or both, is proposed. It may be possible to combine biological and service monitoring for the species because they are directly utilized by humans in subsistence, sport, or commercial fisheries. For example, tagging studies could be designed that provide vital population statistics and exploitation rates. This information may be useful in both assessing recovery and formulating harvest management recommendations. For theses types of studies, ongoing fisheries provide exceptional opportunities to utilize harvest efforts as tag-recapture efforts.

For sockeye salmon populations, which were concluded to be injured at the population level, it may be important to monitor population abundance. If the population is presently protected from exploitation during a recovery period it may be appropriate to conduct detailed spawning ground surveys to assess adult populations. If significant exploitation is occurring or if ongoing population effects are suspected then spawning ground surveys in conjunction with accurate assessment of smolt production may be warranted. Tagging studies (e.g., coded wire tags) may desirable in assessing natural and fishing-related mortality.

Monitoring of spawner escapements for the other species of salmonids may be desirable if the populations are under stress due to harvest or habitat degradation.

Physiological or toxicological analysis may be appropriate on long-lived individuals that may still reflect exposure to the spill or to assess longer term changes in populations due to genetic effects. These analyses would require sacrificing individuals. However, loss of a few individuals should not be a concern in these exploited populations.

6.2.2.4 When to Measure

The timing of sampling will be determined by the parameters selected for monitoring. Consideration must be given to conditions during the sampling period. For example, streams may be at high flow during cutthroat trout spawning periods rendering spawning counts difficult or impossible. Temporal coverage must also be adequate for the questions posed. For example, analysis of spawning populations may need to cover the spawning season sufficiently to enumerate all components of the population.

For physiological or toxicological studies season of sampling is important due to changes that occur due to seasonal variability in food abundance, lipid content, and state of gonad maturation.

6.2.2.5 Data Organization

The data will be dependant upon the questions posed, the sampling methodology, and the analytical methods. Data formats similar to those used by previous researchers are valuable for facilitating comparison of data sets.

The most important data organization issues are data completeness and accuracy. The data set should contain sufficient documentation to allow future users to comprehend the data set. Further, data should be checked and rechecked to ensure all entries are accurate and plausible.

6.2.2.6 How to Analyze

It is highly desirable to analyze the data by robust statistical methods. This will require significant pre-study planning to ensure proper replication and control of the sampled parameters. High variability in biological data often limits the usefulness of the statistical analysis conducted. Sample variability encountered by previous researchers may be invaluable in considering future monitoring activities both in terms of plausibility of statistically valid results and required sample replication. Alternately, limited pre-study sampling may be appropriate to determine population-specific variability during study planning.

In some cases statistical analysis may not be appropriate due to the distribution of the data or the cost of acquiring adequate replication. For example, sufficient replication of genetic tests or residue analysis may not be possible due to budget constraints but these data may still be valuable.

6.2.2.7 How to Interpret

If the study design is statistically rigorous the interpretation of the data should be clear based on the results. However, if particular parameters are difficult to analyze statistically, then more subjective interpretation of the data are necessary. In this case conclusions should be based on the preponderance of evidence rather than individual results.

Greater interpretation of the data may be necessary during hypothesis formulation for future monitoring activities. It is important to differentiate between data interpretation based on statistically valid results and speculation that may be conducted during hypotheses development.

6.2.3 Archaeological Resources

Archeological resources are nonrenewable and do not recover by natural or human-assisted means. Unlike other resource components of the ecosystem, existing prehistoric and historic period sites can never be replaced by natural processes. Archeological resources have a direct link to social, cultural, religious, and scientific values. Because these resources are nonrenewable and represent a link to people's past and future, it is important to ask the question: "Will irretrievable loss (e.g., ethnic heritage value, cultural value) of some archeological sites and artifacts occur if some efforts are not undertaken to restore the injuries?". Specific activities associated with archeological resources that should be considered in developing a sampling design include:

- Direct physical restoration of sites could occur for injuries caused by the oil spill response activities, and looting and vandalism. This activity does not quite meet the definition of recovery used in this conceptual plan, and may best be considered as a restoration activity. This action does not directly involve attempts to test hypotheses about prehistoric and historic period behavior at these sites, but can help to restore the cultural and research value.

Areas of surface disturbance at sites (e.g., looter holes, holes made during clean up activities, ruts from vehicles) could be restored because additional site disturbance (i.e., erosion, looting, and vandalism) can occur if these areas are left in an unrestored condition. Refilling of holes can occur without substantial earthmoving using hand tools and existing soils at the site. Long term monitoring of this type of restoration activity would not be necessary.

- Long-term monitoring could also be considered to monitor the occurrences and rates of vandalism at specific sites. Information from experts indicate that looting and vandalism resulted in the most significant impact to archeological resources.

- Long-term monitoring could be considered to evaluate the effects of oiling on sites because the effect of oiling on chemical components of archeological sites and artifacts is not known.

- Damage assessment data indicate sites that should be considered for restoration actions and that should be considered in integrating archeological resources into the recovery monitoring program.

- As with other injured resources, a reasonable definition or endpoints associated with archeological resources need to be defined. The endpoints for archeological resources are closely related to the nature of the injury.

- Any "recovery", restoration, or long-term monitoring activities for archeological resources must be coordinated with the native groups, other interested parties, and local governments, pursuant to the Archeological Resource Protection Act and the National Historic Preservation Act. Some native corporations require use of their services to access sites on their lands. Local governments should be asked to make recommendations for local sources of services at sites not on corporation lands.

- There are existing database systems for storing archeological data. Archeological data (e.g., location of sites, site descriptions, maps) on federal, native, and state lands is confidential by law. Special contract language needs to be developed for activities associated with archeological resources to protect the data and integrity of the sites.

6.2.4 Intertidal/subtidal

The marine benthic environment exists at the interface of the bottom sediment and the overlying water column. The organisms found in this habitat consist of those that live exclusively in the sediment, those that live exclusively in the water, and those that can make the transition between the water and the sediment. Furthermore, these organisms can range in size from the minute to large. It is important to remember that all of these types of organisms constitute the entire biological assemblage found in a given area.

In sampling marine benthos, it is often logistically impossible or scientifically undesirable to sample for one specific taxon. Rather, the emphasis is on examining the component of the whole assemblage, community, or ecosystem. The sampling and analytical methodology described here is consistent with sampling protocols in use elsewhere (Tetra Tech 1987).
Intertidal and subtidal monitoring programs should concentrate on the numerically dominant or ecologically important taxa and be able to sample them so that predictions and analyses have sufficient statistical power to be meaningful. Nevertheless, it is recognized that indicator or key taxa may not be abundant and that the sampling effort may be altered to address testable hypotheses concerning these rarer taxa.

Preliminary sampling/studies addressing the variations seen in the study area are desirable. These preliminary sampling periods often expose flaws in the sampling or analytical design that are easy to correct prior to the actual study implementation, but which can be difficult to change after the project is fully geared up and functional. Support should be allotted for such preliminary or exploratory sampling.

The optimal sampling design is dependent upon which aspect of the benthos is studied, and upon the habitat being examined. It needs to be emphasized that these components of the benthos are defined by the sampling methods, not by function, ecological interactions, or taxon. Similar sampling methods utilized in different habitats will sample different taxa. As an example, the infauna of embayments open to the ocean will likely be very different from nearby areas of comparable sediments in fjords with a sill across the entrance (Shimek 1990). Regardless of the component of the benthos being examined, after the data are collected,much of the analysis is similar.

No area to be sampled for the benthic assemblages can be assumed to be either spatially or bathymetrically homogeneous. Because of this potential variation, sampling should be at defined stations in the area. Based on preliminary analyses, the data from these stations may be shown to be statistically indistinguishable from station to station. If that is the case, those data may be pooled for subsequent analyses.

The data collected from such sampling would be analyzed with expressed intent of defining and describing the populations of the numerically dominant taxa. The total number and abundance of all collected taxa would be determined, of course; however, the abundant and, presumably, important or target taxa would be the focus of the analyses. Interpretation of the variations seen in these taxa will vary from project to project, depending upon the project design. Nonetheless, the proximate questions generally will involve addressing temporal or spatial changes in population abundances of either particular taxa or groups of taxa.

6.2.4.1 What to Measure

Generally the target taxa will be determined by the questions being asked in the particular monitoring plan. Generally, however, the investigator should concentrate on the numerically dominant and ecologically important taxa. These taxa will need to be sampled so that predictions and analyses have sufficient statistical power to be meaningful. Several descriptive and derived quantitative ecological indices can be used to describe diversity and dominance of the faunal array from each station.

6.2.4.2 Where to Measure

Once the questions that address the benthos have been asked, the decision of where to measure will become obvious. If the questions involve comparisons of data collected over long-term sampling periods, then sampling should be in the same location as previous work. If the questions involve comparisons between current conditions in a spill and reference areas, then paired locations need to be chosen.

6.2.4.3 How to Measure

Assemblages are often measured to discern changes from either the assemblages present at a reference or control area, or differences between the abundances in a sampled area and some pre-defined level of abundance that indicates recovery or restoration. In assessing assemblages of organisms, two measurable factors define many of the observed variations. These parameters are:

- The diversity of the various taxa
- The abundance of those taxa

It must be recognized that changes in these measured factors are themselves due to changes in other more important, but often unknown, parameters that determine the survival and growth of the separate individuals that constitute the populations of each individual taxon.
These ultimate factors may or may not be measurable, but without well-defined experimental projects, their effects will remain uncertain, and will be defined primarily by correlative techniques. All investigators, regulators, and interested readers should be well aware of, and be frequently reminded of, the statistical dictum: "Correlation does not imply causation." The examination of marine benthos by the descriptive-correlative approach, also referred to as the mensurative approach, depends upon the accumulation of a body of observations to support or reject appropriately constructed hypotheses.

These hypotheses must be clearly and precisely phrased to have any validity, and they must be tested with data collected in a manner that insures that confounding hidden factors are minimized. The data need be gathered, analyzed, and interpreted with sufficient awareness of the limitations of the sampling, statistical, and analytical procedures. Clearly, investigators must be cautious about interpreting correlative data without experimental confirmation. However, data collection and sampling plans also must reflect the potential limitations of the procedures involved. Many of those limitations can be addressed by the use of appropriate sampling and statistical methodology.

There are two underlying assumptions to sampling that should be explicitly stated and addressed.

- The first assumption is that the sampling device is physically adequate to sample the populations in question with a minimum of bias. This assumption can be addressed by with a precise definition of the sampled assemblage. Once the assemblage has been defined, an adequate sampling device can be chosen.

- The second assumption is that the sampling plan provides sufficient samples to represent the assemblage. This assumption has to be operationally tested by calculating three related factors:
  - The species-area relationships in the study area
  - The statistical power of the data resulting from the sampling plan
  - The adequacy of the replication

6.2.4.4 When to Measure

The period in which sampling will take place will follow directly from the primary questions being asked, although it may be modified by logistical and seasonal concerns. Annual monitoring will probably suffice for long-term monitoring programs. Nevertheless, seasonal, monthly, or even more frequent sampling periods may be necessary, for example, when questions of reproductive fitness are addressed, the need may exist for sampling gonadal indices over a longer period.

No particular season is likely to provide better data on the organisms than any other. Most Northeastern Pacific benthic infauna show relatively long periods of recruitment and spawning (Strathmann 1987). Consistency from year-to-year, however, will be important in establishing trends.

6.2.4.5 Data Organization

The organization of the sampling and data is dependent upon the questions being asked, the sampling methodology, and the analytical methods. In general, benthic infaunal data consists of a series of station by taxon by abundance matrices. Other data, such as sediment parameters may be also be included in these matrices. Several types of data management systems involving benthic infaunal data presently exist, and most are relatively readily interchangeable.

In general, in addition to the data collected by the sampling effort, some additional data coding will be necessary for the data to be interchangeable between sampling plans or monitoring programs. The most likely candidate code is the National Oceanic Data Center (NODC) coding for individual taxa. In general, two dimensional matrices of variables and observations can be easily manipulated by most spreadsheet software to conform broadly with any other data set, albeit no two monitoring projects will likely collect identical sets of data.

Little attention should be given to requiring the collection of extraneous data. If the data being collected are not pertinent to the project at hand their collection and manipulation can add significantly to the project costs, with no immediate return. This will lead directly to the poor data collection and management and foster questions about the reliability of the required data.

6.2.4.6 How to Analyze

Benthic data are generally voluminous and complex. Analytical methods will vary with the questions being asked. The constraints, assumptions, and inherent strengths and weaknesses of the various analytical methods should be carefully weighed before they are used. The number of valid analytical techniques is too large to simply address here. Individual researchers must be certain of the applicability and adequacy of the techniques they propose.

The utility of inexpensive statistical software programs for microcomputers has put a significant amount of computational power literally at the fingertips of investigators. Unfortunately, many of the statistical subroutines are used without an adequate understanding of the properties of the given statistical test. While is recognized that innovative methods often result in particularly insightful conclusions, investigators are encouraged to be conservative in their utilization of statistical techniques wherever possible. Analysis of variance and t-tests are often the best choices for analyses using parametric statistics, while Friedman Rank-Sum tests and Mann-Whitney U tests would be the corresponding choices as non-parametric tests.

Investigators should make as few assumptions as possible about their data. Particularly, they should not assume that their data are normally distributed; the assumptions inherent in normality can and should be tested. Many valid data transformations exist to allow the data to be analyzed with parametric tests, if those tests are sufficiently more powerful to be desirable in any given situation.

6.2.4.7 How to Interpret

Reference stations are chosen for two reasons. First, to provide indications of overall basin or bay wide changes in the fauna. Numerous bay wide or basin wide changes have been documented in large areas such as Puget Sound in the last twenty years (Nichols 1988), and the occurrence of such changes should be considered in the analyses of the recovery or restoration. Second, reference areas are often used as a benchmark to assess normality of a study area. The study stations and the reference stations are statistically compared, and the results of those comparisons are used in assessing whether or not the study area is "normal."

It should be recognized that it may be difficult or impossible to find true reference stations that are adequate for comparison to the *Exxon Valdez* restoration and recovery stations. Reference stations need to be chosen on the basis of sediment and hydrographic parameters to reflect "normal" or unstressed environments similar to the study areas. Previous work has indicated that shallow water unconsolidated reference areas may be difficult to find. This is primarily due to two factors:

- The hydrographic conditions of any given area are unique and result in a fauna that may not be particularly similar to other sites.

- The assumption must be made that much of the shallow-water benthic habitats in the spill area have been altered, to a greater or lesser degree. Few sites can be found that can be considered *a priori* to be undamaged and thus suitable for providing appropriate monitoring reference stations. Any site must be sampled and analyzed prior to its designation as a reference station. Reference sites are often chosen on the basis of a close match of physical factors and a subjective judgement of normality; the biota are assumed, therefore, to be normal for such an area. Such an assumption without sampling and validation runs the risk of becoming circular: reference areas are chosen because they are assumed to be normal, and then normality of these sites is presumed because they are designated as reference areas.

The first factor precludes use of distant reference stations. The hydrographic conditions in distant areas are likely significantly different from the monitoring stations, and the potential organism groups that may be found might be significantly different due the availability of recruits or unknown physical effects. The second factor means that no or few nearby sites will likely be useful as strict reference area as well.

Rather than try to find a strict reference area for each of the habitats to be sampled, nearby stations may be chosen to provide "background" information about the basic trends in the abundance and composition of the benthos. The background stations will be used to provide an indication of bay or basin-wide changes in the fauna. These background stations will be in located in habitats similar to those being monitored.

Attempts must be made to match the descriptive sediment parameters at the background stations to those of the monitoring stations; however, the background stations may not be chosen specifically to provide the closest possible sediment match to the monitoring stations. These background stations should approximate as many of various monitoring stations' physical parameters as is possible. Although the background stations may not precisely match any particular monitoring station, they should provide a general benchmark in the event of geographically widespread faunal changes.

6.2.5 General Guidance on Sampling Services

Several services provided to the public were also damaged by the *Exxon Valdez* oil spill and, in conjunction with resources, should be considered for monitoring. Use services can be defined as those where individuals directly remove or directly and indirectly use resources. Direct use services include hunting, fishing, hiking, etc. Indirect use services are related to passive or indirect uses (i.e., reading a book about the oil spill area or resources, viewing exhibits in a museum about a resource). Both direct and indirect use services can have a consumptive element (i.e., removing a resource) and a non-consumptive element (i.e., sightseeing, viewing). Although services appropriate for recovery monitoring were identified by the Restoration Working Planning Group and are listed in Section 5, there is an apparent need to make the service and socioeconomic damage assessment data accessible in order to determine measurement parameters important to some services. Some damaged services are directly tied to an economic value (e.g., commercial fishing) of a resource and some are not (e.g., passive uses).

Service-oriented monitoring programs should be integrated with monitoring studies on resource recovery. For example, a specific service-oriented study should be funded if a study is funded on recovery of an injured resource that supports the specific service. In general, probable uses of resources should be paralleled with recovery monitoring activities associated with services for two reasons:

- Services, especially consumptive services, may affect recovery of a particular resource or linkages within the ecosystem, and
- It is important to understand how alleviating or changing the management of a particular service may affect recovery of a resource.

6.2.5.1 Recreation

Recreational services include activities such as sport fishing, sport hunting, boating, kayaking, and camping and hiking. Some of these activities (i.e., sport fishing) have a direct link to some of the injured resources. The nature and extent of damages to recreational services varied by user group and by area of use. Reported changes in use levels and of areas related to avoiding the spill area, reduced wildlife sightings, residual oil, and presence of more people. Changes in people's perceptions of recreational opportunities were also reported. Currently, there are indications that declines in recreational activities reported in 1989 have increased in 1990, but there is no evidence that activity levels have returned to prespill levels.

Factors to consider in selecting monitoring activities associated with recreation and in integrating recreational services into the recovery monitoring program are:

- Recovery monitoring of recreational services should focus on the overlap between the different user groups and the injured resources. For example, if recovery monitoring of harlequin ducks indicates the species is recovering, sport hunting restrictions of the species may be eased, indicating a return or recovery of the sport hunting service.

- Native corporations need to be involved in recovery monitoring activities of recreational resources because they own significant amounts of land used by recreational user groups and they are also major developers of recreational and support activities.

- Recovery or restoration of recreational service is best determined by evaluating increases in use levels (e.g., angler days), and people's perceptions after the oil spill. This implies that there is reasonable information available to develop specific recovery monitoring endpoints.

- Define recovery levels or endpoints for recovery of the specific recreational services need to be defined (See Section 5.3.2.5). One approach to defining a recovery endpoint for recreation is to evaluate existing and pre-spill data to determine what the natural range of variation is for a recreational service and to use the evaluation to identify the variation that will be considered an acceptable endpoint.

- The value of the monitoring activity. Input from economic experts indicate that the value (resulting net economic benefits) of monitoring a particular service, or element thereof, needs to be evaluated by examining the links between the monitoring information to be obtained and habitat changes and population effects.

6.2.5.2 Subsistence

A variety of subsistence resources are used by many residents of Prince William Sound, Kenai Peninsula, lower Cook Inlet, and Kodiak islands. Subsistence resources provide food, resources, and products that are used in daily life and in cultural practices and traditions, and are a means of providing a subsistence-cash economy. Many of the subsistence resources that support a healthy subsistence community are resources that were injured by the oil spill. Pre-spill data is available on subsistence harvests. The two primary endpoints to use to monitor recovery of subsistence harvests are identified in Section 5.3.2.4.
Important factors to consider in planning and implementing a monitoring program for subsistence are:

- As with injured resources and other damaged services a reasonable definition or endpoint of subsistence recovery needs to be defined. Subsistence recovery could be defined as "when the community is harvesting resources (not necessarily the same resources) at a range comparable to pre-spill harvest rates". One approach to defining recovery or a recovery endpoint for subsistence is to evaluate existing harvest data to determine what the natural range of variation is for subsistence harvest, and use the evaluation to identify the variation that will be considered as an acceptable endpoint. A second approach for evaluating the recovery of subsistence harvests is to measure subsistence communities perceptions about contamination of food sources continuing to be dangerous to their health. A third approach that could be integrated with either or both of the other approaches is to integrate recovery monitoring activities of subsistence with specific subsistence resources.

- Involvement of subsistence communities in recovery monitoring. Involvement of subsistence communities in recovery monitoring allows for an opportunity for the communities to take ownership in recovery monitoring activities. For example, cooperative agreements could be established between subsistence communities, the Trustee Council, and an agency or university with expertise and experience in data collection and management. The communities could actually implement a recovery monitoring program with oversight by an agency/university.

- Decide which subsistence communities to monitor. A decision on which subsistence communities to monitor could be determined by evaluating where changes have occurred, and the extent of changes in subsistence harvest that have already been documented, identifying representative communities in the oil spill area, and selecting representative sites within the representative communities.

- Decide what to monitor. Suggestions from interviews with subsistence experts include monitoring levels of participation and shifts in harvest areas, contaminant levels in species that subsistence users depend on, village-wide consumption levels of injured resources that support subsistence, subsistence user perceptions, economic activity of the areas, and market assessments.

- Identify an appropriate method to implement monitoring programs. For example, household interviews are one method that can be used to assess the qualitative measurement of well-being. Interviews could occur initially at all representative sites within representative communities followed by a reduced sampling effort to a few representative sites in representative communities. Using harvest levels as a basis for comparison with pre-spill harvest measurements (taking into account shifts or a new emphasis on certain resources) is also an available method to evaluate levels of participation. Based on input from resource, service, economic, and statistic experts, service parameters that are quantifiable can be treated with statistical methods similar to statistical methods used for natural resources.

- Include subsistence and fishing mortality data (from state and federal catch data) in monitoring recovery activities of particular resources.

6.2.5.3 Commercial Tourism

Commercial tourism is related to passive use values discussed below in Section 6.3.5. Commercial tourism is dependent, in part, on undeveloped wilderness lands and developed lands within the oil spill area. Many areas and resources that support a healthy tourist industry were injured by the oil spill, although the nature and extent of damage varied. There are several types of commercial tourism (i.e., tour ships, day tours, and hunting and fishing charters) that need to be considered and evaluated for the recovery monitoring program. However, the endpoint is very much the same for each of the types of commercial tourism: a return in levels of bookings and reservations to pre-spill levels.

Important factors to consider in planning and implementing monitoring activities for commercial tourist services are:

- Recovery monitoring of tourism should focus on the overlap between the different user groups and the injured resources

- Consideration of the resources that, in part, draw tourists to Alaska

- As with other injured resources and damaged services, a reasonable definition or endpoint for recovery of the specific tourism service needs to be defined. In general, recovery of tourism could be defined as when the levels of use are at a range comparable to pre-spill levels of use. One approach to defining a recovery level or an endpoint for tourism is to evaluate existing data to determine what the natural range of variation is for tourism and use the evaluation to identify the variation that will be considered an acceptable endpoint.

- Decide what to monitor. Suggestions from interviews with service experts include recording the amount of use of particular areas by tourists, monitoring levels of tour boat visitors, and comparing pre- and post-spill ferry passenger data.

- The value of the monitoring activity. Input from economic experts indicate that the value (resulting net economic benefits) of monitoring a particular service, or element thereof, needs to be evaluated by examining the links between the monitoring information to be obtained and habitat changes and population effects. For example, the net benefits of monitoring the commercial salmon fishery could be measured by evaluating how the market would value a resulting change in fish populations and fishery practices. This information could also be used by fisheries managers to achieve more area openings and/or closings and harvest times. Changes in area openings and closures and harvest times can allow for higher quality of fish to be harvested and allow a greater number of fish to be harvested in the short term, long term, or both. However, the probability that the monitoring activity will change future population numbers or quality of harvest would need to be determined. To estimate the value (monetary) gained from the monitoring activity and information, several economic, social, and biological variables would likely need to be collected. These variables could be incorporated into a conceptual model (not conceptual plan) and use the models to forecast the change in fishing industry benefits and fishing industry costs.

6.2.5.4 Commercial Fishing

The commercial fishing industry is the second largest generator of revenue in the state (Exxon Valdez Oil Spill Trustees 1992). Several of the injured resources identified in Section 5 support important commercial fisheries. Recovery monitoring of commercial fishery activities will be important for documenting recovery of the commercial resource and for minimizing impacts to fisheries through fishery management practices. Important factors to consider in planning and implementing a monitoring program for commercial fishing services are:

- Focus monitoring recovery of commercial fishing on the overlap between the different user groups and the injured resources.

- As with other injured resources and damaged services, a reasonable definition or endpoint of commercial fishing recovery needs to be defined. Commercial fishing recovery could be defined as when the commercial harvests are at a range comparable to pre-spill commercial harvest rates. One approach to defining a recovery level or an endpoint for commercial fishing is to evaluate existing commercial catch data to determine what the natural range of variation is for commercial catch harvest and use this evaluation to identify the variation that will be considered an acceptable endpoint.

- Decide which commercial fisheries to monitor. A decision on which commercial fisheries to monitor could be determined by evaluating where changes and the extent of changes in commercial fishing have already been documented, identifying representative fishing communities in the oil spill area, and selecting representative sites within the representative commercial fishing communities.

- Decide what to monitor. Suggestions from interviews with commercial fishing experts include monitoring fishing mortality (from commercial as well as subsistence catch), effects of hatchery production on the service, escapement, economic activity of commercial fishing areas, and market assessments.

- Identifying an appropriate method to implement monitoring programs. For example, the collection of data from fish tickets is one potential method that can provide an inventory and indicator of the health of the fishery.

6.2.5.5 Passive Uses

Passive uses are related to recreational services and tourism and are represented by values that people place on a resource or habitat. Passive users can associate both use and non-use values to a resource. For example, a tourist visiting Pack Creek Bear Preserve may never visit or use McNeil River Preserve, but may value its existence. In addition, non-use values may be derived for a resources existence, or by a desire to pass resources on to the next generation, or intrinsically by deriving some value from the knowledge that the resource remains undisturbed.

Passive use values could also be derived from knowing that there will be option to use the resources in the future.

Passive use values include aesthetic, wilderness, intrinsic, and non-use values. People generally place a high value on knowing that large undeveloped lands provide habitat for fish and wildlife and opportunities for aesthetic enjoyment and appreciation. Important factors to consider in planning and implementing a monitoring program for passive use services are:

- Monitor recovery of passive uses based on the overlap between the different user groups and the injured resources.

- As with other injured resources and damaged services a reasonable definition of a level of or endpoint of passive use recovery needs to be defined. Passive use recovery could be defined based on people's perceptions. One method to determine how society would value efforts to recover lost resources is the application of contingent valuation. Based on information from the Exxon Valdez Oil Spill Contingent Valuation study (via Jeff Hartman, Economist) it is technically feasible to value recovery monitoring activities and restoration activities of passive uses. The application and analysis of contingent valuation would aid the Trustee Council in achieving the goal of maximum and best use of limited funds. The analysis could also provide confidence over which recovery and restoration activities were most highly valued by the public as opposed to making judgements based solely on input from interest groups during the public participation process. Contingent valuation might be applied in the context of valuing tradeoffs between recovery monitoring projects and alternative restoration activities. However, the valuation procedures would not be very effective for defining what the public perception is of a recovery endpoint; it only yields information about the value of the recovery or restoration activity.

- Decide which passive uses to monitor. A decision on which passive uses to monitor could be determined by identifying the non-use values and attempting to quantify those values.

- Identify an appropriate method to implement monitoring programs. For example, well-prepared surveys and interviews could be conducted to determine perceptions of recovery.

6.3 RELATIONSHIP OF THE *EXXON VALDEZ* SPILL MONITORING PLAN TO OTHER MONITORING PROGRAMS

Several programs have been identified that may prove useful to coordinate and/or integrate with the spill monitoring program. Many of these are listed in Table 6, a matrix for identifying elements in common between monitoring programs. Listed below are the programs that may best serve the purpose of this program. In addition to the programs mentioned below, NRDA and restoration science programs should also be reviewed, as discussed earlier, to ensure that methodologies are not reinvented and data is collected in a format that allows the most comparability with the studies previously and/or currently being conducted.

Table 6. Matrix of Exxon Valdez injured resources and elements monitored by other programs.

	Injured at Population Level (Direct Effects)										Injured But No Population Decline										Injured/Damaged Services					
Monitoring Element	SEA OTTER	HARBOR SEAL	COMMON MURRE	MARBLED MURRELET	PIGEON GUILLEMOT	HARLEQUIN DUCK	BLACK OYSTERCATCHER	SOCKEYE SALMON	INTERTIDAL COMMUNITIES	SUBTIDAL COMMUNITIES	KILLER WHALE	RIVER OTTER	BALD EAGLE	CUTTHROAT TROUT	DOLLY VARDEN	PACIFIC HERRING	ROCKFISH	PINK SALMON	ARCHEOLOGICAL SITES/ARTIFACTS	DESIGNATED WILDERNESS AREA	COMMERCIAL FISHING	COMMERCIAL TOURISM	PASSIVE USERS	SUBSISTENCE	RECREATION	Monitoring Programs:
Sediment Chemistry																										EMAP-Near Coastal, Chesapeake Bay Basin, PSAMP, NOAA S&T, Beaufort Sea, Cook Inlet RCAC, Great Lakes, Prince William Sound RCAC
Sediment Toxicity (bioassays)																										EMAP-Near Coastal, PSAMP, Denali National Park & Preserve
Biological Sediment Mixing Depth																										EMAP-Near Coastal
Water Chemistry																										EMAP-Near Coastal, National Surface Water Survey, Chesapeake Bay Basin, PSAMP, Great Lakes, Denali National Park & Preserve
Water Column Toxicity (bioassays)																										EMAP - Near Coastal
Tissue Chemistry (fish and shellfish)																										EMAP-Near Coastal, Chesapeake Bay Basin, PSAMP, NOAA S&T, Cook Inlet RCAC, Great Lakes
Groundwater Chemistry																										Chesapeake Bay Basin, Great Lakes
Submerged Aquatic Vegetation																										EMAP-Near Coastal, Chesapeake Bay Basin, Beaufort Sea (kelp)
Vegetation																										Chesapeake Bay Basin, Denali National Park & Preserve
Habitat Distribution/ Condition																										PSAMP, Denali National Park & Preserve
Benthic: Abundance, Biomass, Species Composition																										EMAP-Near Coastal, Chesapeake Bay Basin, PSAMP, Cook Inlet RCAC, Great Lakes, Denali Park & Preserve
Fish and/or Shellfish: Gross Pathology, Abundance, Species Composition																										EMAP-Near Coastal, National Surface Water Survey, Chesapeake Bay Basin, NOAA S&T, Beaufort Sea (fishery catch data), Cook Inlet RCAC, Great Lakes, NOAA (fisheries), AKF&G (fisheries)
Mussel Watch																										EMAP-Near Coastal, NOAA S&T, Beaufort Sea, Cook Inlet RCAC, Prince William Sound RCAC
Zooplankton																										National Surface Water Survey, Chesapeake Bay Basin, Great Lakes
Phytoplankton																										Chesapeake Bay Basin, Great Lakes
Bacteria																										Chesapeake Bay Basin
Birds: Water-, Land-Based																										Chesapeake Bay Basin, PSAMP, Beaufort Sea (oldsquaw, common eider), Great Lakes (comorants), Denali National Park & Preserve
Reptiles, Amphibians																										Chesapeake Bay Basin
Mammal(s): Abundance, Tissue																										PSAMP, Beaufort Sea (bowhead whale, ringed seal), Denali National Park & Preserve (small land mammals), FWS (seabirds, sea otter, boat bird surveys), NMFS (harbor seal, sea lion)

6.3.1 Resource Monitoring Programs

6.3.1.1 Alaskan Monitoring Programs

- Cook Inlet Regional Citizens Advisory Council monitoring program
- Prince William Sound Regional Citizens Advisory Council monitoring program
- Agency resource management programs such as Alaska Fish and Game programs (e.g., salmon, subsistence, harbor seals, salmon catch, sea otter food habits and reproduction, mussel bed contamination, etc.)
- Oil Spill Response Institute proposed monitoring program
- Oil Spill Health Task Force Group
- National Park Service social indicator study, intertidal and coastal programs
- Coastal Marine Institute (Univ. of Alaska and Minerals Management Service)
- Any programs developed with the criminal settlement funds

6.3.1.2 Federal Programs in Alaska

- Agency resource management programs such as U.S. Fish and Wildlife Service programs (e.g., bird and sea otter monitoring)
- NOAA Status and Trends/Mussel Watch/Benthic Surveillance program
- Denali National Park and Preserve
- National Marine Fisheries Service harbor seal program and Marine Mammal Protection Act surveys
- NOAA restoration studies, and hydrocarbon monitoring program and database
- Any programs developed with the criminal settlement funds

6.3.1.3 Private or Other Programs

- Sea World killer whale and humpback studies
- British Columbia killer whale monitoring program (Federal)

- Global programs such as WOCE and Tropical Ocean and Global Atmosphere
- Coastal Regional Monitoring Act/Program (Regional Marine Research Program)
- U.S. Geological Service remote sensing
- NOAA weather service
- U.S. Fish and Wildlife Service (USFWS) Environmental Studies Program
- Joint Canadian, Russian, and USFWS program on seabirds

6.3.1.4 Future Programs With Which to Coordinate

- NOAA Status and Trends program
- U.S. National Park Service coastal program
- EPA's Environmental Monitoring and Assessment Program (EMAP) - Near Coastal
- U.S. National Park Service archaeological program

6.3.1.5 Monitoring Programs to Learn From

- Puget Sound Ambient Monitoring Program
- Beaufort Environmental Monitoring Program
- Chesapeake Bay
- Great Lakes
- Santa Monica Bay estuary/restoration program (Los Angeles Regional Water Control Board)
- OCSEAP studies by OCS and MMS

6.3.2 <u>Service Monitoring Programs</u>

- U.S. Forest Service recreational use surveys
- Mineral Management Services survey of subsistence use in coastal areas
- ADF&G subsistence survey and chemical contamination data
- ADF&G commercial fisheries (e.g., salmon escapement surveys and herring spawn deposition)
- Restoration Planning Work Group survey of perceptions and use levels
- Oil Spill Health Task Force Group

- National Park Service social indicator study
- Any programs developed with the criminal settlement funds

7. IMPLEMENTATION AND MANAGEMENT OF MONITORING PROGRAM

7.1 IMPLEMENTATION AND MANAGEMENT

Implementation of the monitoring program will occur during Phase 3. The overall monitoring effort should be managed by a single contractor or team of contractors (a prime contractor with subcontractors), as should the coordination of the centralized data library. The contractor(s) should work with an advisory team that consists of the various user groups, including principal investigators, peer reviewers, public, and agency staff, as well as Restoration Team members.

Another option for managing the monitoring is to utilize a system similar to that used by the Puget Sound Ambient Monitoring Program (PSAMP). Management of the monitoring would be handled by a Monitoring Management Committee (MMC) made up of 15 to 30 individuals representing the various user groups. Formation of an MMC is similar to the recommendation that a contractor manage the overall monitoring effort, except for who is the lead party. With the development of an MMC, an agency would most likely take the lead. Due to the political nature of the *Exxon Valdez* program, we recommend that the Trustee Council utilize a contractor versus placing the responsibility with an agency. This adds some objectivity to the program.

Continuing with the system utilized by PSAMP, the MMC or other managing body, would be responsible for completion of Phase 2 of the program. Public and peer review would be received prior to implementation (as discussed further below), with the final outcome of Phase 2 resulting in recommendations to the Trustee Council for implementation of the monitoring program (Phase 3). Management recommendations for Phase 3 might also stem from a PSAMP-type process, with the development of an institutional structure to coordinate and manage the program. Basically this consists of a steering committee formed of the agencies and institutions implementing the monitoring program. The advantage of this is that the parties conducting the monitoring have an active role in managing it, and thus is presumably understanding, coordinating and utilizing the results of the program.

The format for the individual monitoring elements (Phase 2) can either be contracted under one contract or under individual contracts for specific resources and services, or monitoring type. However, all contractors must agree to comply with a set of guidelines in order to be awarded the contract. Whether or not the monitoring is contracted as a single project or as multiple projects also depends on the funding available for monitoring. It may be most cost efficient to have a single contractor perform the work since the QA/QC, overhead expenses, etc., would all be covered by a single element. Technically this could be better as well, because the sampling effort and techniques would remain consistent throughout the program. (However, information may be lost if the resource or service expertise is not applied.)

Management of the program consists of coordinating not only the implementation of the program but the reevaluation phases of the program, including peer review. The most certain way to ensure that the data are collected, analyzed, and presented in a scientific and meaningful manner is to remove as much bias as possible from the selection and funding processes, while directing the projects with a set of requirements and guidelines. A competitive bid process is recommended, utilizing peer reviewers from the proposal to the final award stage. To achieve the goals and objectives established in the conceptual monitoring plan, monitoring activities will require effective, well coordinated management. There are many competing interests for the settlement funds. There are also numerous competing objectives and goals outside those generally agreed to in the conceptual monitoring plan. Thus, effective, well coordinated management is essential to leading monitoring activities in a manner that will attain the overall monitoring goal.

What are the elements of effective management? Although this question has many different answers, there are several components of management that will help to ensure that it is effectiveness:

- Make decisions in a logical manner.
- Direct activities toward established goals.
- Involve interested parties.
- Make decisions on a timely basis.
- Communicate decisions immediately to the involved parties.

To accomplish effective management it often requires authorities to delegate more responsibility than they wish. It requires all interested parties to take substantial risks that their interest may not receive the highest priority.

At a more elemental level, a process or mechanism, such as a detailed schedule with trigger dates, can be developed for each of the project funding areas, such as monitoring, damage assessment, and restoration. These can then be overlaid to provide the Trustee Council with an overall restoration schedule from which to plan. Feedback from the public, in the form of review by the Public Advisory Group, or public comments on draft project documents, can be used by the Trustee Council to determine the priorities of the public.

Feedback from principal investigators and peer reviewers can be used by the Restoration Team and Trustee Council to determine priorities and trigger points for scientific concerns. For instance, in order for a restoration activity to be underway in summer 1993, the principal investigator may need six months' notice to allow for securing logistical support. Therefore their monitoring proposal must be reviewed and a funding decision made with at least six months' advance notice. Priority should be given to schedules that are calendar-driven or otherwise inflexible so as not to lose information. Using the example above, priority would be given to an advanced review of the proposed study so that a field season of data collection is not lost.

There may be trigger dates that overlap between project areas or time lines that are impossible for the Trustee Council to meet. These should be negotiated at the onset of planning. In situations such as these, the Trustee Council should utilize outside expertise to prioritize and/or reschedule activities and/or if possible, delegate some of their responsibilities.

The schedule described above, should be continually revised and updated, but should be relatively stable on a quarterly basis, to allow the Trustee Council to plan ahead. In addition to trigger dates, dates can be backed into the calendar to allow the Trustee Council advance warning of an upcoming event. Again, using the example above, if the Trustee Council needs a two-month period to send a proposal out for expert technical review prior to its own review of the proposal, this date can be backed in, as well as the review period necessary for the Trustee Council to complete its review.

7.1.1 Peer Review Panel

A panel of peer-reviewers could be selected to review all stages of program design and implementation. The reviewers should review and grade all proposed projects, following guidelines developed by the Trustee Council or utilizing a format similar to that used by a well accepted funding agency, such as the National Science Foundation (NSF). A similar peer-review process should be used for all project renewals, and for review of draft and final reports. Projects should be fully funded by the restoration funds with no subsidiary funding by a private or state agency in order to eliminate any potential conflicts of interest. This does not preclude agency scientist from bidding on monitoring elements; however , the contract award and work conducted must be discreet from other activities conducted by the agency. Matching or assisting funding might be allowed if the funding agency was independent such as the NSF or the National Institutes of Health.

A peer-review panel could be selected by using lists of NSF reviewers, or from the National Academy of Sciences. Additional reviews could be done through the mail, as per NSF. The panel should be relatively large, six to twenty members, and should reflect all types of biases. If the proposals are ranked on merit, the resulting reviews will be relatively unbiased as the extreme views will balance each other. Personal bias is unavoidable and unremovable, but the relatively unbiased selection of meritorious projects, coupled with adequate QA/QC procedures should foster the development of projects whose results will reflect the unbiased nature of the analyses.

Because it is expected that some of the monitoring activities may continue for several years, it may be useful to have a rotating review panel, with the rotations staggered like that of many board of directors. For instance, each reviewer can serve on the panel for only two years at a time. The first terms would be of a staggered length (one to three years) such that all the peer reviewers would not be lost at the same time.

In addition to resource and service experts as peer reviewers, a team of statisticians and modelers to review program design(s) is recommended. It would be most useful if this team could act as a resource available to all projects teams, as well as in the review capacity to ensure comparability of the programs.

This team of biostatisticians can also be involved in the proposal review stage and with recommending upgrades or changes to a program. For instance, a strong proposal may come in but have inadequate sample replication or statistical applications. Rather than disqualify the proposal, changes can be recommended, and a best and final technical and cost proposal requested.

7.1.2 Data Dissemination

All of the monitoring data should be kept in a central repository or library, accessible by a computerized system linking the available databases. How and who can utilize this system will be a decision of the Trustee Council, but oversight of the repository will be the responsibility of whomever is managing the monitoring program.

It is important that the monitoring results be made known to the public. This may take the form of summary fact sheets, summaries of activities in the Restoration Work Plan, or in another form selected by the Trustee Council.

The method for data dissemination to the different users is not an objective of the monitoring program, rather that data be accessible and in a format that can be readily utilized by scientists, resource managers, investigators and other interested parties.

7.1.3 Avoiding Duplication of Effort

As discussed above, integration and/or coordination with other programs is essential to avoid duplication of effort. In order to avoid this duplication it is essential to coordinate monitoring efforts between studies, both those that are funded with settlement funds, and those that are not.

To facilitate the coordination between programs it would be useful to do the following:

- Develop a computerized and hard copy table that identifies ongoing routine agency monitoring activities for resources and services affected by the oil spill or that occur within the oil spill area. (Incorporating GIS for maps may be useful for this purpose).

- Communicate with state and federal resource agencies to follow changes in routine agency monitoring activities.

7.2 REQUEST FOR PROPOSAL AND CONTRACT LANGUAGE

Developing contracts to ensure timely performance requires incentives for completing tasks on schedule and disincentives for tardiness. Incentives for completing tasks on time could include financial bonuses or some type of preferred status in selection for future rounds of project work. Disincentives could include the loss of money through financial penalties, or exclusion from consideration for any further project work. For example, standard contract language could require the contractor to inform the client within a certain number of days if there was a problem that would affect schedule. However, early notification to the client must be tied to some type of penalty for being late with deliverables without notification.

If the funding for a project comes from public sources, an additional means of encouraging timely performance might be a regular (monthly/quarterly) public review of project contracts. This could take the form of a public meeting where contractors had to explain where they were in the project and why they are not meeting the schedule. An alternative would be a regular display ad in local newspapers indicating who is responsible for different projects and whether they were on schedule.

Examples of a few contract scenarios that use the basic ideas of incentives and penalties are described below:

Assumptions:

- Lump sum type contract
- $100,000 total contract value
- $20,000 mobilization costs

7.2.1 Payment Tied to Deliverables/Schedule

This form of contract would incorporate a number of deliverables such as reports or milestones in the monitoring process and payment would come after satisfactory completion of those identified tasks.

Example: If the project had four equal cost deliverables or milestones, the contractor would receive $20,000 up front to get started and would receive the next $20,000 payment upon completion of Task 1. Upon completion of Task 2, another $20,000 would be paid.

The contractor does not receive money for the next phase until they produce the deliverable. Thus, money is withheld until project completion (Task 4).

Pros/cons: This approach would work well with sequential tasks. The contractor would not have funds to work on Task 2 until they successfully completed Task 1. This assumes both public and private groups would want a relatively steady stream of funding to avoid extreme staffing fluctuations.

This idea would not work as well if the tasks have to occur concurrently, or if the contractor has to be paid up-front. If the tasks occur simultaneously, then staffing levels would be committed and there would not be an incentive to complete tasks in a timely manner.

The same problem would exist if the contractor is a public agency and needs total funding up-front before beginning work. In this case, they have already received the money and there would not be a clear incentive for them to meet the schedule without including a penalty. A penalty, such as withholding the final 10 to 15 percent of the contract total until receipt and acceptance of the final product may be sufficient penalty.

7.2.2 Percent Reduction in Contract Total Value

This concept includes a penalty clause that reduces the total value of the contract based on lack of timeliness. A percentage of the total value of the contract is established as a penalty and is withheld for each day/week the product is late.

Example: If the project has four equal cost deliverables, the contractor receives $20,000 up front to get started. If they are one week late delivering a report and the penalty is calculated at one percent per week, then the penalty is $1,000 ($100,000*.1). This penalty could be taken out of each specific task (i.e. payment of $19,000 for completion of Task 1), or it could be withheld from the last payment for completing Task 4. If the contractor was one week late on Task 1, met the schedule in Tasks 2 and 3, and was one week late in Task 4, the final payment would be $18,000.

This approach could also be used to create an incentive. If the product or milestone is reached early, a reward could be established that would provide additional money or some other benefit to the contractor.

7.2.3 Incentive for Continuing Project Involvement

This approach assumes there will be a built in incentive for public/private groups to continue their involvement with the project. For example, an agency or private entity will want to be associated with doing the long-term monitoring of sea otters.

Example: If a contractor successfully meets their deadlines/milestones in Year 1, they are automatically given first opportunity to do similar work in Year 2. If they miss a deadline, then the work in Year 2 is out for competitive bid and the contractor runs the risk of losing the work. If the contractor decides to bid on Year 2 after losing automatic "rehire" rights, they will have to explain to the satisfaction of the Trustees their lack of performance in Year 1.

Depending upon the details of the specific monitoring project, a combination of these concepts could be developed to include both incentives and penalties in the same contract.

7.2.4 Schedule for Deliverables, Performance Criteria and Proposal Ranking

- Schedule for deliverables so if problems arise client must be informed in advance and presented with solution

- Performance criteria for meeting QA/QC requirements, standard protocols, compatible data

Again, an attachment could be used in a contract document to identify what specific protocols will be used, what criteria will be used to assess compliance with QA/QC requirements, and the data format to be used for compatibility.

- Factors for proposal review and requirements for ranking/rating proposals

This isn't a contractual issue and should be handled on a technical basis. A copy of the contract with payment provisions could be included in the Request for Proposal (RFP). The contractor could be asked to comment on any potential problems they see with the contract format. This information could be helpful in determining the successful candidate.

8. RECOMMENDATIONS

A summary of the recommendations stemming from development of the conceptual plan are presented below. The recommendations are presented as eleven elements. Further elaboration is provided in the previous sections of the report.

- The process or mechanisms we recommend to assist the Trustee Council in determining monitoring priorities consists of the following:

  Consensus building. Involve all user groups in the restoration/recovery process, including the public, agencies, and scientists. This is recommended

  - To ensure that those affected by the spill are involved in recovery of the spill area and use of the settlement funds

  - To ensure buy-in by as many people as possible in order to gain and hold their support

  - To ensure resource and service management agencies as well as scientists are involved in order that their data needs are understood

  Development of recovery endpoints. Recovery endpoints should be further refined. Recovery endpoints developed in Phase 1 need to be further refined in Phase 2 with the input from at least three experts on each resource and service. Recovery endpoints are necessary to construct testable hypotheses.

  Application of criteria. Prioritization of the resources and services should initially be conducted through application of the criteria defined in this report. The rank of each criterion for each resource and service needs to be developed from the input of at least three experts on each resource and service. This information can then be used to prioritize the resources and services for monitoring, and reviewed along with costs for each monitoring element that will be obtained in Phase 2.

  Conceptual model development. Conceptual models for each resource and service to be monitored should be developed to better understand the biological, chemical and physical processes and interactions affecting the resource or service, as well as the social, cultural and economic factors. Conceptual models will aid in focusing the monitoring on testable hypotheses, and in interpreting the results of the monitoring. The linkage matrix and recovery endpoints contained herein can be used as a start to development of the conceptual models.

- The Trustee Council has developed a list of injured resources and damaged services. Injured resources with direct (population level) effects and indirect (sublethal) effects are considered for monitoring. Those that should be monitored should be based foremost on those that through monitoring, provide the most information on recovery of the ecosystem effected by the *Exxon Valdez* oil spill, and those that may provide information on the effects of future perturbations. The resources and services should be prioritized according to the criteria presented herein, most of which are scientifically based, but some of which take into account cultural and socioeconomic concerns. Once this prioritization is complete, the monitoring elements for the resources and services should be evaluated (not prioritized) based on costs. This process may lead to some changes in monitoring priorities if, for instance, a resource monitoring program has the highest priority based on scientific merit, but costs two million dollars, and the next three studies of high technical priority cost $250,000 each. If only one million dollars is available, then the three $250,000 programs may take priority over the two million dollar project. Linkages between resources and services should also be evaluated, along with the potential effect of physical and chemical processes.

- Whenever possible, the Trustee Council should take advantage of existing, proven methodologies, either from NRDA and restoration studies, or from other monitoring programs. This will help ensure a successful program and provide the most comparable and long-term database for evaluation.

- Whenever possible, (whenever the goals are similar or can be mutually addressed), the Trustee Council should coordinate or integrate it's program with other monitoring programs, including resource agency programs. To aid this, completion, updating, and expansion of the matrix developed in this plan should be accomplished in Phase 2. The value is in producing a comparable, long-term database.

- The monitoring program should be managed by an independent contractor or agency that does not have political interests in the direction the program takes. The independent contractor or agency should have access to an advisory group consisting of representatives of the users groups, including principal investigators.

- The program should have continuous feedback and review steps to ensure that the needs and objectives of the program are being met. Review should take the form of independent review by a rotating pool of experts in the resources and services being monitored, as well as by biostatisticians, modelers and perhaps economists. Peer review should be implemented at all stages of the program including proposals, sampling design, results, and final products.

- The monitoring elements that will be outlined in Phase 2 should go through a competitive bid process to ensure that the best programs and expertise are utilized.

- The data should be centralized in a database network or library, that is user friendly and designed to meet the needs of the various users.

- Guidelines (both at the proposal and contract level) should be developed to standardize the quality of the deliverables received, and to facilitate use of the various databases generated.

- Penalties, such as withholding payment or going out to bid if deliverables are not submitted on schedule or meeting the QA/QC guidelines, should be included in the contract phase to ensure that schedules, cost and quality of deliverables is achieved.

- Projects awarded should be ensured funding throughout the time period necessary to document recovery, as long as the objectives of the monitoring program and contract requirements are being met. Because natural recovery is estimated to be between four and 120 years for the resources and services identified by the Trustee Council, an endowment, or other source of long-term funding will be necessary and is recommended.

9. REFERENCES

Ballou, T., S. C. Hess, R. Dodge, A. H. Knap and T. D. Sleeter. 1989. Effects of untreated and chemically dispersed oil on tropical marine communities: a long-term field experiment. Proceedings of International Oil Spill conference. American Petroleum Institute. 447-454.

Chan, G. L. 1977. The five-year recruitment of marine life after the 1971 San Francisco oil spill. Proceedings of the International Oil Spill Conference. American Petroleum Institute. 549-558.

Clark, R. C., B. G. Patten, and E. DeNike. 1978. Observations of a cold-water intertidal community after five years of low-level, persistent oil spill from the *General M. C. Miegs*. J. Fish. Res. Board Can. 35:754-765.

Conan, G. 1982. The long-term effects of the *Amoco Cadiz* oil spill. Phil. Trans. R. Soc. Lond. 297:323-333.

Cretney, W. J., C. S. Wong, D. R. Green, and C. A. Bawden. 1978. Long-term fate of a heavy oil in a spill-contaminated B. C. coastal bay. J. Fish. Res. Board Can. 35:521-527.

Dauvin, J-C. and F. Gentil. 1990. Conditions of the peracarid populations of subtidal communities in northern Brittany ten years after the *Amoco Cadiz* oil spill. Mar. Pollution Bull. 21:123-130.

Dauvin, J-C. 1987. Evolution a long terme (1978-1986) des populations d'amphipodes des sables fins de la Pierre Noire (Baie de Morlaix, Manche Occidentale) apres la catastrophe de l'Amoco Cadiz. Mar. Environ. Res. 21:247-273.

Elmgren, R., S. Hansson, U. Larsson, B. Sundelin, and P. D. Boehm. 1983. The *Tsesis* oil spill: acute and long-term impact on the benthos. Marine Biology. 73:51-65.

Environment Canada. 1982. Review and Evaluation of Adaptive Environmental Assessment and Management. Prepared by ESSA Environmental and Social Systems Analysts Ltd. Vancouver, British Columbia, Canada. 116p.

Exxon Valdez Oil Spill Restoration Office. 1993. Draft *Exxon Valdez* oil spill restoration plan: summary of alternatives for public comment. Anchorage, Alaska.

Green, R.H. 1979. Sampling Design and Statistical Methods for Environmental Biologists. John Wiley & Sons. New York, N.Y. 257p.

Gulliksen, B. and J. P. Taasen. 1982. Effect of an oil spill in Spitzbergen in 1978. Marine Pollution Bulletin. 13:96-98.

Jacobs. R. 1980. Effects of the *Amoco Cadiz* oil spill on the sea grass community at Roscoff with special reference to the benthic infauna. Marine Ecology Progress Series. 2:207-212.

Linden, O., R. Elmgren, and P. Boehm. 1979. The *Tsesis* oil spill: its impact on the coastal ecosystem of the Baltic Sea. Ambio 8:244-253.

National Research Council and Marine Board. 1990. Managing Troubled Waters: The Role of Marine Environmental Monitoring. National Academy Press. Washington, D.C. 125p.

Nichols, F. H. 1988. Long-term changes in a deep Puget Sound benthic community: local or basin-wide. *In:* Proceedings, First annual meeting on Puget sound research. Puget Sound Water Quality Authority. Seattle, Washington. pp. 65-71.

Notini, M. 1978. Long-term effects of an oil spill on *Fucus* macrofauna in a small Baltic bay. J. Fish. Res. Board Can. 35:745-753.

Poole, R. W. 1974. An introduction to quantitative ecology. McGraw-Hill Book Company. New York, New York. 532 p.

Shimek, R. L. 1990. Diet and habitat utilization in a Northeastern Pacific Ocean scaphopod assemblage. American Malacological Bulletin. 7:147-169.

Strathmann, M. F. 1987. Reproduction and development of marine invertebrates of the Northern Pacific coast. University of Washington Press. Seattle. 670 pp.

Teal, J. M. and R. W. Howarth. 1984. Oil spill studies: a review of ecological effects. Environmental Management 8:27-44.

Tetra Tech. 1987. Recommended protocols for sampling and analyzing subtidal benthic macroinvertebrate assemblages in Puget Sound. Prepared for U.S. Environmental Protection Agency, Region 10, Office of Puget Sound, Seattle, Washington.

Westman, W.E. 1985. Ecology, Impact Assessment, and Environmental Planning. John Wiley & Sons. New York, N.Y. 532p.

APPENDIX A

QUESTIONNAIRE FOR DEVELOPMENT OF PRELIMINARY DRAFT OF CONCEPTUAL PLAN

Questionnaire No. 1

QUESTIONS FOR INTERVIEWS WITH "EXPERTS"

1. What are your expectations of the monitoring program?
   - of the conceptual plan?

2. What do you think the focus of the monitoring program should be?
   - what do you want to achieve?
   - what are your concerns?
   - are there any regulatory requirements you are aware of?

3. What do you believe are the primary goals of the monitoring?

4. How would you determine if recovery is adequate?

5. Do you have an opinion on the resources and services that should be monitored?
   - any prioritization?
   - why do you believe these resources/services are important?

6. What regional monitoring programs should this monitoring program be integrated with?

7. How best do you believe the data can be used as decision-making tools?

8. Is there anything important to the conceptual monitoring plan that we have not asked you about?

QUESTIONNAIRE NO. 2

QUESTIONS FOR INTERVIEWS WITH EXPERTS

1. What do you think a conceptual monitoring plan should accomplish?

Objectives

2. After reviewing the list of objectives which, if any, should be changed and are there other objectives we should consider?

Strategies

3. After reviewing the list of strategies which, if any, should be changed and are there other strategies we should consider?

Criteria

4. Can you tell us how you think the stated criteria should help direct the decision making process for restoration monitoring?

5. Is there other information or questions that we should consider as criteria for decision making in defining restoration monitoring activities?

6. What, if any, priorities or weightings would you give to the listed criteria as part of an objective decision making process?

7. What criteria are the most useful or least useful given your area of expertise?

Other Questions

8. Are there special innovative forms of statistical design and analysis which you are currently involved with that may apply to a given resource or service?

9. What kinds of qualitative analyses, if any, should be applied effectively to a monitoring program if a resource or service is not amenable to a quantitative analysis?

10. What damage assessment databases are you aware of that we should consider?

    What are the strengths/weaknesses of these databases?

11. What characteristics do you think are most important in developing an integrated database?

12. What if any monitoring programs that you are aware of do you think the restoration monitoring should be coordinated with?

13. How would you recommend incorporating services into this conceptual monitoring plan?

14. Are you familiar with specific survey instruments which effectively assess impacts/restoration on injured services like recreation, subsistence, aesthetics, etc.? If so, are there any kinds of qualitative data collection that can be of use to this monitoring plan?

15. Which surveys of services (e.g., recreation, subsistence, aesthetics) provided by natural resources contain elements that would best serve the purposes of the conceptual monitoring plan?

What are these elements?

16. Is there anything else that you think is important to include in the conceptual monitoring plan that you would like to comment on?

APPENDIX B

LIST OF INTERVIEW PARTICIPANTS

INDIVIDUALS PARTICIPATING IN INTERVIEWS

RESOURCE SERVICE REPRESENTED	CONTACT	RESTORATION TEAM	RESTORATION PLANNING WORK GROUP	PEER REVIEWER	PRINCIPAL INVESTIGATOR
OVERVIEW, MARINE ECOLOGY					
	Bob Spies			X	
	Don Boesch			X	
	Pete Peterson, PhD			X	
	Gail Irvine, PhD				X
	Bruce Wright			X	
	Doug Wolfe, PhD			X	
subtidal communities	Stan Rice, PhD			X	
coastal communities	Ray Highsmith, PhD				X
benthic communities	Steven Jewett, PhD				X
	Al Mearns, PhD				X
MARINE MAMMALS					
pinnipeds	Don Siniff, PhD			X	
killer whales	John Ford, PhD			X	
sea otters	Jim Bodkin				X
river otters	Jim Faro				X
humpback whale/killer whales	Marilyn Dahlheim, PhD				X

RESOURCE SERVICE REPRESENTED	CONTACT	RESTORATION TEAM	RESTORATION PLANNING WORK GROUP	PEER REVIEWER	PRINCIPAL INVESTIGATOR
harbor seals	Kathy Frost				X
FISHERIES					
	Ray Hilborn			X	
salmon	Phil Mundy, PhD			X	
	Joe Sullivan, PhD				X
pink/chum salmon	Sam Sharr				X
sockeye salmon	Dana Schmidt				X
BIRDS					
seabirds	Vern Byrd - *see river otters too*				X
seabirds	George Hunt, PhD			X	
sea ducks	Sam Patten				X
bird restoration	Dennis Heineman, PhD			X	
bird toxicology	Michael Fry, PhD			X	
ARCHAEOLOGY					
	Martin McAllister, PhD			X	
	Doug Reger				X

RESOURCE SERVICE REPRESENTED	CONTACT	RESTORATION TEAM	RESTORATION PLANNING WORK GROUP	PEER REVIEWER	PRINCIPAL INVESTIGATOR
RECREATION					
	Jim Richardson			X	
	Jon Isaacs			X	
SUBSISTENCE					
	Jim Fall, PhD				X
RESOURCE ECONOMICS					
commercial fisheries	Lewis Queirolo				X
STATISTICS/POPULATION BIOLOGY					
population biology	Lee Eberhardt, PhD			X	
statistics	Doug Robson, PhD			X	
DECISION ANALYSIS					
	Ken Reckhow, PhD			X	
	James Ruttenber, PhD			X	
MICROBIOLOGY					
	Joan Braddock, PhD				X
MARINE CHEMISTRY					
	Jeffrey Short				X
GIS					
	Art Weiner				X
TOXICOLOGIST					
	John Stegeman			X	
RESTORATION PLAN WORK GROUP					
	Karen Klinge		X		
	Mark Kuwada		X		

RESOURCE SERVICE REPRESENTED	CONTACT	RESTORATION TEAM	RESTORATION PLANNING WORK GROUP	PEER REVIEWER	PRINCIPAL INVESTIGATOR
RESTORATION TEAM					
	Mark Broderson	X			
	Pamela Bergmann	X			
	J. Jerome Montague	X			
	Byron Morris	X			
	Ken Rice	X			

APPENDIX C

LIST OF WORKSHOP PARTICIPANTS

ATTENDANCE AT WORKSHOP ____________

RESOURCE SERVICE REPRESENTED	PERSON	AGENCY/ASSOCIATION
OVERVIEW, MARINE ECOLOGY		
	Bob Spies	Applied Marine Sciences (API)
	Don Boesch	University of Maryland
	Gail Irvine	U.S. National Park Service
MARINE MAMMALS		
sea otters	Jim Bodkin	U.S. Fish and Wildlife Service
river otters	Jim Faro	Alaska Department of Fish and Game
river otters	Vern Byrd - *see sea birds too*	U.S. Fish and Wildlife Service
FISHERIES		
salmon	Phil Mundy	Peer Review - API
BIRDS		
seabirds	Vern Byrd - *see river otters too*	U.S. Fish and Wildlife Service
bird restoration	Dennis Heineman	
seabirds	Michael Fry	University of California, Davis
seabirds	Sam Patten	Alaska Dept. of Fish and Game
ARCHAEOLOGY		
	Martin McAllister	Archaeological Rsc. Invest.
RECREATION		
	Jim Richardson	Resource Econ., Inc.
	Jon Isaacs	Jon Issacs and Associates
ECONOMICS		
	Jeff Hartman	Alaska Department of Fish and Game
STATISTICS/POPULATION BIOLOGY		
	Doug Robson	Peer Review - API

RESOURCE SERVICE REPRESENTED	PERSON	AGENCY/ASSOCIATION
MARINE CHEMISTRY		
	Jeffrey Short	NOAA/NMFS Auke Bay Fisheries Lab
COASTAL COMMUNITIES		
	Ray Highsmith	University of Alaska, Fairbanks
RESTORATION TEAM		
	Mark Broderson	Alaska Dept. of Environmental Conservation
	Pamela Bergmann	Department of Interior
	Marty Rutherford	Alaska Department of Natural Resources
	J. Jerome Montague	Alaska Department of Fish and Game
	Byron Morris	NOAA/NMFS
	Dave Gibbons	U.S. Department of Agriculture
	Ken Rice	U.S. Forest Service
RESTORATION PLANNING WORK GROUP		
	John Strand	NOAA/NMFS/Restoration Office
	Carol Gorbis	U.S. Fish and Wildlife Service
	Mark Kuwada	Alaska Department of Fish and Game
	Ray Thompson	U.S. Forest Service
	Karen Klinge	U.S. Forest Service
	Chris Swenson	Alaska Department of Fish and Game
	Veronica Gilbert	Department of Natural Resources
REGIONAL CITIZENS ADVISORY COUNCIL		
	Dennis Randa	Cook Inlet RCAC
	Jim Day	Cook Inlet RCAC
	Shelli Vacca	Prince William Sound RCAC
OTHERS		
	Joe Sullin	Alaska Dept. of Fish and Game

APPENDIX D

TEN STEPS TO A STRONG MONITORING PROGRAM

TEN STEPS TO A STRONG MONITORING PROGRAM[1]

CLEAR GUIDANCE: as to how data are to be used and type of decisions to be made.

GOALS: establish scientifically, technologically, logistically and financially achievable goals.

DECISION INTEGRATION: decision points and feedback loops should be clearly established and integrated into decision process prior to data collection.

AUTHORITY & CONTROL: explain, then define where they reside and provide local controls compatible with program controls and objectives.

PARTICIPANT COMMUNICATION: identify communication channels among participants and ensure they are both functional and interconnected.

INTEGRATE NEEDS: for regulation, data acquisition and management of local, state and federal agencies to optimize use of available resources.

PUBLIC & SCIENTIFIC INVOLVEMENT: establish mechanisms to ensure these groups are participants early and often.

COMMUNICATION: establish mechanisms to ensure conclusions are communicated to both decision makers and the public in terms they can understand and act upon.

REVIEW: include mechanisms for periodic review and re-direction when results or information justifies a change.

MANAGEMENT: identify in advance, actions to be taken in response to expected and unexpected results.

---

1 Restated from Natural Research Council. 1990. Managing Troubled Waters, National Academy Press, Washington D.C.